Explaining Chaos

Discovering Lorenz System's Extreme Sensitivity to Initial Conditions

Tuomo Lahtela

Master Thesis

Supervisor: Jani Lukkarinen

Department of Mathematics and Statistics

University of Helsinki

28.09.2019

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Matemaattis-luonnontieteellinen	Matematiikan ja tilastotieteen laitos

Tekijä — Författare — Author
Tuomo Lahtela

Työn nimi — Arbetets titel — Title

Explaining Chaos: Discovering Lorenz System's Extreme Sensitivity to Initial Conditions

Oppiaine — Läroämne — Subject
Matematiikan aineenopettaja

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu -tutkielma	Syyskuu 2019	62 s.

Tiivistelmä — Referat — Abstract

Tutkielmassa pyritään kertomaan lyhyt tarinanomainen esitys kaaosteoriasta. Esitys tarkastelee ja selittää kaaosteoriaan olennaisesti liittyviä käsitteitä kuten deterministisyys ja alkuarvoherkkyys. Oleellisesti tutkielma kertoo kaaosteorian isänäkin pidetyn Edward Norton Lorenzin tarinan maailman ensimmäisen kaaottisen systeemin löytymisestä ja sitä myötä alkuarvoherkkyyden käsitteen syntymisestä. Tutkielman tarkoitus on näyttää aiheeseen perehtymättömällekin lukijalle mistä kaaosteoriassa on kyse sekä miksi se on merkityksellistä.

Johdantokappaleen jälkeinen luku on jaettu kolmeen osaan, jotka käsittelevät kaaosteorian asettumista tieteenhistorian jatkumoon, käsitettä deterministisyys ja tapahtumaketjua, jonka seurauksena Edward Norton Lorenz teki tieteellisen löytönsä. Kolmas luku selittää helposti lähestyttävin esimerkein käsitteen alkuarvoherkkyys, joka tunnetaan paremmin myös nimellä perhosvaikutus.

Neljännessä luvussa esitellään dynaamisen systeemin käsite selittäen sen olennaisuuden maailman tapahtumia esitettäessä ja formuloitaessa matemaattisesti. Luvussa paneututaan myös matemaattisemmin käsitteeseen deterministisyys. Luku 5 tutustuttaa lukijan tutkielman tärkeimpien lukujen 6,7 ja 8 ymmärtämiseen tarvittaviin matemaattisiin menetelmiin kuten Taylorin sarjateoriaan useammassa ulottuvuudessa, Jacobiaaniin sekä linearisaatioon.

Luku 6 esittelee tutkielman pääaiheen, Lorenz-systeemin, määritellen sen matemaattisesti sekä kuvaillen sen ominaisuuksia yksinkertaista sääsysteemiä mallintavana systeeminä. Luvussa käydään läpi myös Lorenz-systeemin ymmärtäminen vektorikentän käsitteen kautta ja systeemin ratkaisupolun geometrinen representaatio. Luvussa tutkitaan myös onko Lorenz-systeemi alkuarvoherkkä kaikkialla lähtöavaruudessa. Luvun lopussa näytetään myös erittäin kuvaannollisti kuinka Lorenz-systeemin käyttäytymisen ennustaminen on käytännössä mahdotonta.

Luvussa 7 Lorenz-systeemi osoitetaan alkuarvoherkäksi seuraamalla systeemin aikakehitystä. Oleellisesti kahden alkuarvoiltaan miltein identtisten ratojen välistä etäisyyttä mittaavan vektorin aikakehitystä seuraamalla näytetään, että radat erkanevat toisistaan erittäin nopeasti. Viimeinen luku esittää lyhyen yleisanalyysin alkuarvoherkkien systeemien aikakehityksestä. Luku esittelee myös kaaosteoriaan olennaisesti liittyvät käsitteet attraktori ja outo attraktori. Luvun loppuun on vielä tiivistetty tutkielman otsikkoa kunnioittaen kaaoksen suppea selitys listaten kolme päätekijää kaaoksen käsitettä matemaattisesti määrittämään.

Avainsanat — Nyckelord — Keywords
Kaaosteoria, Lorenzin systeemi, alkuarvoherkkyys, perhosvaikutus, outo attraktori

Säilytyspaikka — Förvaringsställe — Where deposited
Kumpulan tiedekirjasto

Muita tietoja — Övriga uppgifter — Additional information

Contents

Chapter 1

Introduction

In a chaotic fashion, this mathematics thesis starts not with a quote from a mathematician or a chaos theorist but a psychoanalyst Carl Jung: "In all chaos there is cosmos, in all disorder there is order [Jun69, p. 32]." Jung's approach to chaos was not too mathematical, but the following 60 pages are more so; they tell a story about chaos theory. Respectively, this thesis is a humble attempt to explain the mathematical concept of chaos by acquainting the reader with the *secret order in all disorder*-part of Jung's quote, whereas the *cosmos*-part is unfortunately given less attention.

Chaos theory is occasionally in danger of being overtaxed by its association with everything that can be even superficially related to the concept of chaos. Unfortunately, a sometimes extravagant popularization through the media is also contributing to this danger. But at the same time this popularization is also an important opportunity to free areas of mathematics from their intellectual ghetto [PJS04, p. 9]. Along these lines this thesis endeavours to distinguish chaos from concepts often mistaken as such, while still trying to keep also the less mathematically oriented reader aboard. The story is essentially wrapped around the Lorenz system and its extreme dependence on initial conditions.

The journey starts in Chapter 2 with a historical glance over chaos theory. We look at how scientific understanding of the world slowly turned into an attempt to mathematize it. We briefly flash back to how 17th century fellows Newton and Leibniz prepared the basis for modeling the laws of movement from which the term determinism, an essential concept of this thesis, was furthermore bolstered by Laplace in early 19th century. After studying what Heisenberg's 1927 uncertainty principle has to do with all this we explore with further detail the term determinism and see why it earned its significance in science vocabulary, and what separates it from the world of probability. We make also the bed for later discussion about chaos emerging from determinism. The last section of Chapter

2 is dedicated for Edward Norton Lorenz's discovery of *initial value sensitivity* which is an essential property when defining chaos.

In Chapter 3 we show in black on white what *initial value sensitivity* means. Two tables are provided to illustrate the numerical evidence on how initial value sensitivity truly exists, and how it emerges even from very simple systems. Furthermore the metaphor *butterfly effect* is introduced and explained.

As the main focus of the following pages is the Lorenz system and as it is very much a dynamical system, Chapter 4 seats to answer the question *"What's a dynamical system?"* in detail. After the general interpretation of what could be taken as a dynamical system, we proceed to a more mathematically rigorous definition. The context of dynamical systems allowing so, we define the term determinism better. Additionally, differential equations and planar systems are included into the picture as well.

Chapter 5 presents some fundamental mathematical methods which are used in the following chapters. Taylor series is introduced and proved to function in higher dimensions, a concept of Jacobian Matrix is briefly explained and lastly a method of linearization is familiriarized to the reader. The less mathematically oriented reader might want to consider simply skimming through the Taylor series part, as it might give an impression of a more difficult read, and focus on the Jacobian Matrix and linearization.

The keen reader can see Appendix A for all the relevant proofs, lemmas and background theory and Appendix B for all the MATLAB code standing behind the mathematics of Chapters 6, 7 and 8. The main focus of this thesis, the Lorenz system, is introduced in Chapter 6. The original derivation of the system is described and all of the variables and parameters are explained. Naturally, the mathematical formulation of the Lorenz system

$$\begin{aligned} \frac{\partial x}{\partial t} &= a(y-x) \\ \frac{\partial y}{\partial t} &= x(c-z)-y \\ \frac{\partial z}{\partial t} &= xy-bz. \end{aligned} \tag{1.1}$$

is presented together with the parameter values a, b and c producing chaotic behavior. To meet the less mathematically oriented reader, we take a look how the Lorenz system can be interpreted as a vector field and how each of the vectors with their components can be illustrated geometrically. We also look at how the solution path geometrically lands on $\mathbb{R}^3$ when we fix an initial condition.

In Chapter 7 we study the Lorenz system's sensitivity to initial conditions. This is done by following the evolution of two initially close-by trajectories. In short, we show that the deviance vector measuring the distance between two trajectories grows exponentially over time and prediction of the system breaks down. Because the exponential growth coefficient turns out to be of matrix form, we are destined to estimate the deviance vector length through theory of Jacobian matrices and their eigenvalues. In Section 7.4 we briefly study the topic of *"Is the Lorenz system sensitive to initial conditions everywhere?"*. The topic is approached by mapping the maximum eigenvalues of Jacobian matrix in $\mathbb{R}^3$ and loosely locating the regions where the Lorenz system would exhibit sensitivity to initial conditions.

In Chapter 8 we do our best to analyze the Lorenz system through means other than just accurate future prediction. We take a look what extreme initial value dependency means when geometrically viewing the system as a whole. We also introduce two very essential terms, *attractor* and *strange attractor*. With these terms, we finally reach our concluding Section 8.5 of Chapter 8 in which we try to, in a very compact manner and to the best and humblest of our knowledge, explain chaos.

Chapter 2

Chaos: A Brief Survey

2.1 Historical Glance

In the ages of early human history, most of the natural events must have seemed to be of pure chaos in the less mathematical meaning of the word. The lens of scientific understanding through which the world is today viewed was very blurry. But with increasing speed the natural sciences developed. The development can be interpreted as shrinking of the area where chaos, the lack of human understanding, reigned. The governing laws were recognized for more and more nature's phenomena. Simultaneously with natural sciences, the mathematical formulation developed. This soon meant that the understanding of the nature of a phenomena would include the appropriate mathematization of them. This concurrence of harnessing of natural phenomena and mathematization of them contributed to the illusion that it was only a matter of time before chaos would be completely extinct.

No more than three hundred years ago a milestone accomplishment of colossal and accelerating effect was made. Through the universal mathematical ideas of calculus Sir Isaac Newton (1643-1727) and Gottfried Wilhelm Freiherr von Leibniz (1646-1716) provided the basis for modeling the laws of movement. Be it the planets, the development of populations, the spread of sound through gases, the conduction of heat in media, the interaction of magnetism and electricity, or even the course of weather events, the fundamental framework how to model them was discovered. The findings of Newton and Leibniz carried with them the false belief that the terms determinism and predictability were equivalent.

Another name worth mentioning regarding the history of determinism is Pierre-Simon Laplace. In 1814 Laplace defined an imaginary demon carrying his surname as "A con-

sciousness great enough to know the exact locations and velocities of all the objects in the universe at the present instant, as well as all the forces. Thus there could be no secrets from this consciousness. It could calculate anything about the past or the future from the laws of cause and effect." As determinism became more and more grounded in calculus, thus in mathematical formulation of the world also, the Laplace's demon grew to be the symbol of it.

What the deterministic credo is saying is that the universe is comparable to the ordered running of a astoundingly precise clock, in which the present state of things is, on the one hand, the cause of its future state, and, on the other hand, the consequence of its prior state. Present, past and future are hemmed in by causal relationships which often take form of mathematical formulation. According to deterministic view, the problem of an exact prediction was only a matter of the difficulty of recording all the relevant data. The Newtonian era was characterized by the deterministic credo. For the natural sciences this era came to an end at the latest via the findings of Werner Heisenberg in the 1920's.

Heisenberg, in his 1927 proclamation of the uncertainty principle, wrote "In the strict formulation of the causality law - When we know the present precisely, we can calculate the future - it is not the final clause, but rather the premise, that is false. We cannot know the present in all its determining details. Therefore, all perception is a selection from an abundance of possibilities and a limitation of future possibilities... Because all experiments are subject to the laws of quantum mechanics, and thereby also to the uncertainty principle, the invalidity of the causality law is definitely established through quantum mechanics." Classical determinism in its exceptious strictness had to be buried – a turning point of enormous importance [PJS04, p. 11-12].

The history of numerical weather forecasting depicts profoundly the undiminished belief in a deterministic world as in reality, Heisenberg's uncertainty principle did not at all mean the end of determinism. The absolute mathematical precision presupposed by Laplace is not physically possible as minute imprecision is always present. Despite this ubiquitous imprecision, scientists all over the world still kept, and in some cases keep, believing that from approximately the same causes follow approximately the same effects – in nature as well as in any good experiment. And this is often the case, a fact worth pinpointing here, especially over short time spans. If this would not be the case we would not be able to establish any natural laws.

But this apparently plausible assumption relating to approximate causes and approximate effect is not universally correct - over long time periods it does not describe the typical course of natural processes or dynamical systems modeling them. Lucky for our universal

mathematical understanding of the world an American mathematician and meteorologist Edward Lorenz discovered this assumption might be incorrect [PJS04, p. 13-14]. But before learning more about Lorenz's discovery we try to grasp the concept of *determinism* and explain its stand in the natural sciences and chaos theory.

2.2 Deterministic Chaos

Despite its name, chaos theory falls into the category of the last decades' greatest accomplishments of natural sciences. Very few developments in natural sciences have awakened such public interest. Chaos theory can even be seen as breaking and revolutionizing images of reality. But what is it in chaos theory that captivates the human interest? Or what do the supposed revolutions in the image of reality consist of and how? Posing and answering these questions leads us immediately to the fundamentals of philosophy of nature and origins of natural sciences.

The main maxim of science, its ability to relate and present cause and effect, takes its most obvious form in the natural sciences. For example on the basis of the laws of gravitation eclipses, appearances of comets and other astronomical events can be predicted thousands of years in advance. This is obviously not the case with all the natural phenomena. For instance, although the everyday evolution of the atmosphere surrenders to the laws of physics just as much as the movements of the planets do, weather prediction is still to this day exceedingly problematic.

It's a question of language too. We often speak of the unpredictable aspects of weather or other deterministic events in the same fashion as we were talking about rolling dice or letting an air balloon loose to observe its irregular path. As no clear relation between cause and effect, such phenomena are taken as random elements. Yet it is, or was, erroneously conceived that precise predictability could, in principle be achieved. Before chaos theory it was assumed that all that's needed is to gather and process greater quantities of more precise information. In the case of weather prediction this would mean denser networks of weather stations and more capable computers to analyze the data.

The conclusions of chaos theory have altered this viewpoint. Seemingly random behavior can be generated by simple deterministic systems with only a few elements. That random behavior should be emphasized here as it is truly fundamental; gathering more information does not make it disappear. This fundamental randomness deserves its name. And the name is chaos [PJS04, p. 9-10].

The magic separating chaos theory from probability theory is that chaos is deterministic. Chaos is generated by fixed rules, often even simple rules, which do not themselves involve any elements of randomness. We even speak of deterministic chaos. In essence the future is completely determined by the past and the past is completely determined by the future but when minute errors of measurement enter into calculations they amplify to an extent that even though the behavior is predictable in the short term, it is unpredictable over the long term.

When considering chaos theory, the discovery of such unpredictable behavior is one of the most important achievements. But almost as important are the methodologies which came to be designated for a specific scientific evaluation of the presence of chaotic behavior in mathematical models as well as in real phenomena. These methodologies enable, at least in principle, to estimate the predictability horizon of a system. Predictability horizon is the mathematical, physical, or time parameter setting a limit within which predictability is ideally possible and beyond which we will never be able to predict properly.

The predictability horizon for example in weather forecasting is never more than approximately two or three weeks. What this signifies is that no matter how many more weather stations are included in the observation, no matter how much more rigorously weather data are collected and analyzed, the weather, despite being deterministic, can never be predicted with any degree of numerical accuracy beyond this time horizon [PJS04, p. 10-11].

2.3 Edward Norton Lorenz

How did deterministic chaos came to be found then? In 1959 a mathematician and professor of meteorology at the Massachusetts Institute of Technology, set himself a challenge that would shake the foundations of science of prediction. Edward Lorenz challenged himself to find a simple, nonlinear mathematical model of a weather system that would generate non-periodic weather patterns. In 1959 strong doubt was the dominant reaction towards existence of non-periodic solutions to equations without any random elements. Nonetheless, Lorenz fearlessly promised to deliver a paper titled “The statistical prediction of solutions of dynamic equations”.

Fortunately, Lorenz had the necessary, and requisite, tool for beating his challenge: a primitive personal computer. This very rudimentary computer nonetheless enabled Lorenz to explore the nature of the weather patterns produced by simple deterministic models. And with some adjusting of parameters Lorenz soon discovered the non-periodic solutions

he was searching for. Somehow, a dynamical system without explicit uncertainty was producing apparently unpredictable motion. Lorenz had, against quite the odds, found the system he needed to write his paper.

Continuing to experiment with the system Lorenz found something else as well. By a minor accident he had run two calculations with almost but not quite identical initial values resulting to completely different solution patterns. On further examination, Lorenz noted that the difference between the two runs did not arise immediately when the simulation started. The discrepancy crept in gradually, disturbing the least significant digits first, then higher order digits, and finally giving way to a completely different results.

What had happened, how come the results were so completely different? In systems with friction, or viscosity in the case of fluid, an initial offset usually decays rapidly and is soon lost entirely. In Lorenz model viscosity was deterministic, in other words built into it, making it more surprising that the initial offset developed to such great extents. Still, he was certain about the accuracy of his calculations and accepted the fact that the non-periodic weather patterns he had calculated were highly unstable. What he had discovered was concept which eventually made its way also to topic of the thesis you are reading: *initial value sensitivity.*

Perhaps more importantly, Lorenz soon realized that this sensitivity to initial values, in other words instability, might just be the explanation for why weather is so difficult to predict. Meteorological prediction rely on deterministic simulations and measurements of pressure, wind speed, temperature, etc., taken at various points as initial values of the variables of interest. There is and will always be a limit for measurement accuracy. Hence any weather prediction based on model as unstable as Lorenz's would have serious limitations [Kau11, p. 145-146].

First the skeptics discounted Lorenz's work because his weather model was very simple, involving a dozen variables rather than the thousands or perhaps millions required to assure accuracy. But where Lorenz's instinct as a mathematician kicked in was when he followed his gut feeling of trying to show that it is possible to find instable nature in system much simpler than the one he had discovered in 1959. So in 1961 instead of hearing his skeptics complaints about a too low a number of prediction variables he lowered the number even more finding a set of equations with just three variables, usually taken as x, y and z, that display the same kind of unstable, non-periodic behavior as his 1959 weather model [Kau11, p. 146].

In the scale of the big picture Lorenz showed that natural laws, and especially deter-

minism, do not exclude the possibility of unpredictability. By showing that determinism and predictability are not equal he truly revolutionized the image of reality. He also demonstrated how unpredictable effects can be produced in very simple systems and that chaos and order, referring here to the causality principle, can be found together within the same system. "One of the lessons coming out of chaos theory is that the validity of causality principle is narrowed by the uncertainty principle from one end as well as by the intrinsic instability properties of the underlying natural laws from the other end." as phrased by Peitgen, Jürgens and Saupe in their comprehensive work *Chaos and fractals - new frontiers of science* [PJS04, p. 13-14].

The main focus of this thesis is to study the initial value sensitivity of Lorenz's famous system of three variables and to tell what it means to not to be able to predict the behavior of a deterministic system. But before diving deeper into Lorenz's system we explore some fundamental theory related to chaos as it is mathematically understood. All this is done to make the beauty of the Lorenz system apprehensible.

Chapter 3

Initial Value Sensitivity

One way to see the concept of a system is to picture something that takes an input, usually a number value, and out of this input it gives an output, usually a number value. A system based prediction, very shortly explained, is presuming that if a certain input value p_0 gives a certain output value p_1 then a input value close to p_0 would give an output value close to output value p_1. A system that is sensitive to initial conditions affectively gives significantly differing output values for significantly similar input values. Another name for such behavior is dependence on initial conditions. Put the name aside, such behavior is a distinguishing characteristic of chaotic behavior. A dynamical system that displays sensitive dependence on initial conditions will produce significantly different solutions for two specifications of initial states that are initially very close together. In point of fact, given any initial condition, there is another condition close to it that will diverge from it by some required distance, when observation period is long enough.

Lorenz spelled out the consequences of his discovery of initial value sensitivity as follows: "It implies that two states differing by imperceptible amounts may eventually evolve into two considerably different states. If, then, there is any error whatever in observing the present state – and is any real system such errors seem inevitable – and acceptable prediction of an instantaneous state in the distant future may well be impossible". Systems that display sensitive dependence on initial conditions exhibit what Lorenz labeled the *butterfly effect*. This nowadays disturbingly famous dictum he used to allegorize how even tiny flaps of a butterfly can make a difference when predicting the weather [Kel94, p. 12].

Lorenz and his discoveries are introduced and explained more in detail in the subsequent chapters . In this chapter we represent examples of data sets to introduce the concept of sensitivity to initial conditions. The main purpose of the following examples is to show how even simple deterministic systems can produce chaos.

3.1 Data Sets To Illustrate Chaos

The first experiment iterates the quadratic expression $p+rp(1-p)$ for constant $r = 3$ and initial value $p_0 = 0.01$ using two calculators produced by different manufacturers. The difference between the calculators is that the other, Casio, is restricted to ten decimals while the other's, HP's, accuracy is twelve decimals. We conjecture that over time the calculators' *result suggestions* will differ from each other.

3.1.1 Two Calculators

Table of the iterations on both calculators is provided below in Table 3.1. Neither one of the calculators can exactly represent the iterations three and higher as the second iterate need eight decimals and would therefore yield to sixteen decimals in the third iterate. Hence, there are cut-off errors which can not be avoided but which at first look don't seem to matter much. This at least seem to be the case when we look at the iterations four and five where the results of both calculators agree in ten decimal places [PJS04, p. 46-49].

Evaluations	Casio	HP
1	0.0397	0.0397
2	0.15407173	0.15407173
3	0.5450726260	0.545072626044
4	1.288978001	1.28897800119
5	0.1715191421	0.171519142100
10	0.7229143012	0.722914301711
15	1.270261775	1.27026178116
20	0.5965292447	0.596528770927
25	1.315587846	1.31558435183
30	0.3742092321	0.374647695060
35	0.9233215064	0.908845072341
40	0.0021143643	0.143971503996
45	1.219763115	1.23060086551
50	0.0036616295	0.225758993390

Table 3.1: Two slightly different calculators iterating the same simple quadratic expression [PJS04, p. 49].

When comparing the calculators we might apt to trust the HP more as it works with

two decimal places higher accuracy. One might say we accept the HP answer for the final iterate and conclude the Casio's result to being wrong. This is a common bias. The Casio's failure results from a principal mathematical problem and the HP is subject to the same problem. We can not assume that the two decimal advantage in calculation accuracy would make the problem disappear. Actually, the only thing we can say for sure is that one of the two calculators is totally wrong. It is also very likely that both calculators are off despite the fact that the deterministic process of logistic map, $p + rp(1 - p)$, is very simple. This sensational effect is the inevitable consequence of finite accuracy arithmetic which every machine from pocket calculator to super-computers are forced to use. This means that no matter how powerful a computer one uses the effect can not be avoided.

The minuscule differences in the two calculators accumulate so rapidly that their predictive power fade away. But things can get even wilder as we are to see in the following experiment.

3.1.2 One Calculator

This time we run our example of the quadratic dynamic law, $p+rp(1-p)$, with exactly the same values as before $r = 3$ and $p_0 = 0.01$ but just on one calculator. Two comparative runs are completed with the higher-accuracy calculator HP. The only thing changed is the way the calculator evaluates the quadratic expression. How much difference can it make if the expression is evaluated as $p + rp(1 - p)$ instead of $(1 + r)p - rp^2$?

Table 3.2 indicates quite a big difference. It really goes against one's intuition of how mathematical equations work. The error creeps in already at the 12th iterate as the 11th decimal differ from the other [PJS04, p. 51].

Evaluations	$p + rp(1-p)$	$(1+r)p - rp^2$
1	0.0397	0.0397
2	0.15407173	0.15407173
3	0.5450726260	0.5450726260
4	1.288978001	1.288978001
5	0.1715191421	0.1715191421
10	0.7229143012	0.7229143012
11	1.323841944	1.323841944
12	0.03769529734	0.03769529724
13	0.146518383	0.1465183826
14	0.5216706225	0.5216706212
15	1.270261775	1.270261774
20	0.5965292447	0.5965293261
25	1.315587846	1.315588447
30	0.3742092321	0.3741338572
35	0.9233215064	0.9257966719
40	0.0021143643	0.0144387553
45	1.219763115	0.0497855318

Table 3.2: Two different implementations of the same quadratic law on the same calculator [PJS04, p. 52].

3.2 Butterflies Matter

If not the first experiment then the second should convince the reader that sensitivity to initial conditions is very much present and alive. With the finite accuracy computing our escape from the effects chaos is limited, perhaps forever. Predictability breaks down eventually.

Limited computing machines, also known as computers, really shape our understanding of reality. The fact that initial value sensitivity and thereby chaos have entered upon the scientific stage is essentially an achievement made possible by new tools provided to science by computers. As notable progress in the deeper understanding of chaos theory is being achieved it is becoming clear that chaos actually follows certain very stable patters. One could say that order can be found from chaos. Strangely, this profound understanding of chaos was and is again discovered by means of computers, which in different circumstances seem so vulnerable to chaos [PJS04, p. 52].

Chapter 4

Dynamical Systems

The Lorenz system, the main focus of this thesis, is essentially a dynamical system. This is why in this chapter we explain what is a dynamical system and introduce its mathematical definition.

4.1 What Is a Dynamical System?

Any mechanism that evolves deterministically in time can be taken as a dynamical system. The simplest examples, or at least simplest mathematical representations, of dynamical systems can be found in dynamics. Anything from pendulums to Solar Systems and from atoms reacting with each other to weather systems spanning over the whole globe can be treated as dynamical systems. We should note that when treating biological and economic systems as dynamical systems it is less clear when we are dealing with determinism as random elements often complicate the picture. In physics context, *equations of motion* describe a physical system's behavior as a set of mathematical functions in terms of dynamical variables. Usually this means that spatial coordinates are presented as functions of time. Main interest of study of dynamical systems is evolution functions that describe the state of a system as a function of time satisfying the equations of motion of the system.

The simplest type of evolution is of course stationary, where the state remains constant in time. The next most studied type of evolutions are periodic evolutions where after a fixed period, the system always returns to the same state as the system repeats precisely the previous period. Stationary and periodic evolutions are very regular and predictable.

The term *prediction* is controversial but generally in the context of dynamical systems it is based on matching with past geometrical and numerical observations. When steering away from stationary and periodic evolutions one meets not so regular and predictable

evolution already in quite simple systems. When unpredictability can be brought about by deterministically defining a system, we shall speak of chaotic behavior [BT09, p. 1-2].

4.2 Determinism Defined

When constructing a definition for dynamical systems an essential property is determinism, a property we discussed already in Chapter 2. By definition within deterministic system the present state, meaning both the position and the velocity, determines all future states. Furthermore, the whole past can be reconstructed from the present state. The key concept here is that of state, coined as position and velocity in the case of dynamical systems. All possible states together form the state space. For a function $\varphi = \varphi(t)$ that satisfies the equation of motion, we call the points $(\varphi(t), \varphi'(t))$ of the state space an evolution. The curve $t \to (\varphi(t), \varphi'(t))$ can also be called evolution.

Next we aim to express the unique determination of each future state by the present one, in other words determinism, in terms of a map. If $(\varphi(t), \varphi'(t))$ is the present state (at $t = 0$) and $t > 0$ in the instant future, then we denote the state at time t by $\Phi(\varphi, \varphi', t)$. This expresses the fact that the present state determines all future states. So, if $\varphi(t)$ satisfies the equation of motion, then we have

$$\Phi(\varphi(0), \varphi'(0), t) = (\varphi(t), \varphi'(t)).$$

The map

$$\mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^2$$

constructed in this way, is called the evolution operator of the system. One should note that such an evolution operator can 'reconstruct the past'.

Generally dynamical systems are defined as structures consisting of state space, also known as phase space. Here state space is indicated by M. Therefore and evolution operator takes the form of

$$\Phi : M \times \mathbb{R} \to M. \tag{4.1}$$

Now we shall develop a set of fundamental properties for evolution operators of dynamical systems. The first property to be discussed is that for any $x \in M$ necessarily

$$\Phi(x, 0) = x. \tag{4.2}$$

(4.2) means that when a state evolves during a time interval of length 0, the state remains unchanged. The second property we mention is

$$\Phi(\Phi(x, t_1), t_2) = \Phi(x, t_1 + t_2). \tag{4.3}$$

What (4.3) is essentially saying is that if $t \to \varphi(t)$ satisfies the equation of motion, this also holds for $t \to \Phi(t + t_1)$, for any constant t_1.

4.3 When Differential Equations Enter the Picture

Interestingly enough systems of first order differential equations are almost identical with just defined dynamical systems. In fact, if for the state space M we take a finite-dimensional vector space and Φ of class at least C^2, then the C^1-map $f : M \to M$ can be defined by

$$f(x) = \frac{\partial \Phi}{\partial t}(x, 0).$$

If, and only if, curve $t \to x(t)$ is a solution of the differential equation

$$x'(t) = f(x(t)), \tag{4.4}$$

it is an evolution of the dynamical system defined by Φ. The evolution operator $\Phi : M \times \mathbb{R} \to M$ in this case is often called the 'flow' of the ordinary differential equation (4.4) [BT09, p. 15-17].

4.4 Planar Systems

Dynamical systems are often defined over one single independent variable, commonly thought as time t. This custom emanates likely from human interest to see how things change over time. Depending on the system observed its state might constitute of multiple components. And as at least natural systems tend to be more complex than simple, this is often the case. Thus we introduce the topic of *systems of differential equations.* A system of differential equations is a collection of n interrelated differential equations of the form

$$\begin{aligned} x_1' &= f_1(x_1, x_2, ..., x_n) \\ x_2' &= f_2(x_1, x_2, ..., x_n) \\ &\vdots \\ x_n' &= f_n(x_1, x_2, ..., x_n). \end{aligned}$$

The functions f_j are real-valued functions of the n variables $x_1, x_2, \ldots, x_n$. We take that the f_j are C^∞ functions. This means firstly that the partial derivatives of all orders of the f_j exist and secondly that they are continuous.

To simplify notation we use vector notation $X = (x_1, \ldots, x_n)$. Thus our a system of differential equations takes a more concise form as

$$X' = F(X),$$

where

$$F(X) = (f_1(x_1, x_2, \ldots, x_n), \ldots, f_n(x_1, x_2, \ldots, x_n)).$$

Likewise a function of the form $X(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ represent a solution for system alike. In regards of existence and uniqueness of such solution the Existence and Uniqueness Theorem is presented in Appendix A.1. The system of equations is called autonomous if none of the f_j depends on t and likewise the system is called non-autonomous if some of the f_j depends on t, when the system becomes $X = F(t, X)$. This thesis is concerned only with autonomous systems [HDS74, p. 21-22].

Chapter 5

Mathematical Methods: Taylor Series in Higher Dimensions, Jacobian Matrix and Linearization

In this chapter we present mathematical methods that are essential to understand before proceeding to the following chapters. First we introduce the all familiar Taylor series but in higher dimensions making it a little knottier to approach. Second we introduce the concept of Jacobian matrix very shortly and third we familiarize the reader with the concept of linearization.

5.1 Taylor Series in Higher Dimensions

Dedicating a section of its own for Taylor series serves us in two ways. First it provides a thorough tool to approximate function's local behavior in multiple dimensions. Second, it presents an estimation of how good of an approximation Taylor series actually is for function's local behavior. Both manners play admittedly fundamental roles when investigating initial value sensitivity which by definition pertain high accuracy approximations. Section 5.1 follows closely G.B Holland's presentation [Hol, p. 2-4].

5.1.1 Multi-Index Notation

It is helpful to introduce some new notation when generalizing Taylor's theorem. A multi-index is an n-tuple of nonnegative integers. We use Greek letters α or β to denote multi-indices

$$\alpha = (\alpha_1, \alpha_2, ..., \alpha_n), \quad \beta = (\beta_1, \beta_2, ..., \beta_n), \quad \text{where} \quad \alpha_j, \beta_j \in \{0, 1, 2, ...\}.$$

We also will be using the following definitions

$$|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_n$$

$$\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$$

$$\boldsymbol{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

$$\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}.$$

The order of α is defined as $|\alpha| = \alpha_1 + \cdots + \alpha_n$. Consequently the order of α is the same as the order of ∂^α as a partial derivative or the order of $\boldsymbol{x}^\alpha$ as a monomial. If f is a function of class C^k, the kth order partial derivatives of f can be written simply as ∂^α with $|\alpha| = k$. For example with $n = 3$, $\boldsymbol{x} = (x, y, z)$ and $k = 2$ we have multiple different αs. Examples of such are presented as

$$\alpha^{(0,2,0)} f = \frac{\partial^2 f}{\partial y^2} \quad \text{and} \quad \alpha^{(1,0,1)} f = \frac{\partial^2 f}{\partial x \partial z}.$$

Just as the notation $\boldsymbol{x}^\alpha$ implies, multi-indices are helpful when writing not only derivatives but also polynomials in several variables. In the following a generalization of the binomial theorem is presented [Hol, p. 1].

5.1.2 The Multinomial Theorem

Theorem 5.1.1. *(The Multinomial Theorem). For any $\boldsymbol{x} = (x_1, x_2, \ldots x_n) \in \mathbb{R}^n$ and any positive integer k the following holds*

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \boldsymbol{x}^\alpha.$$

Proof. The case $n = 2$ is just the binomial theorem

$$(x_1 + x_2)^k = \sum_{j=0}^{k} \frac{k!}{j!(k-j)!} x_1^j x_2^{k-j} = \sum_{\alpha_1+\alpha_2=k} \frac{k!}{\alpha_1!\alpha_2!} x_1^{\alpha_1} x_2^{\alpha_2} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \boldsymbol{x}^\alpha,$$

where $\alpha_1 = j, \alpha_2 = k - j$ and $\alpha = (\alpha_1, \alpha_2)$.

The general case follows by induction on n. Suppose the result is true for $n < N$ and $\boldsymbol{x} = (x_1, x_2, \dots x_N)$. By using the result for $n = 2$ and then the result for $n = N - 1$, we obtain

$$\begin{aligned} (x_1 + x_2 + \ldots + x_N)^k &= [(x_1 + x_2 + \ldots + x_{N-1}) + x_N]^k \\ &= \sum_{i+j=k} \frac{k!}{i!j!} (x_1 + x_2 + \ldots + x_{N-1})^i x_N^j \\ &= \sum_{i+j=k} \frac{k!}{i!j!} \sum_{|\beta|=i} \frac{i!}{\beta!} \tilde{x}^\beta x_N^j, \end{aligned} \tag{5.1}$$

where $\beta = (\beta_1, \beta_2, \dots, \beta_{N-1})$ and $\tilde{x} = (x_1, x_2, \dots, x_{N-1})$.

To conclude we set $\alpha = (\beta_1, \beta_2, \dots, \beta_{N-1}, j)$, so that $\beta!j! = \alpha!$ and $\tilde{x}^\beta x_N^j = \boldsymbol{x}^\alpha$. Observing that α runs over all multi-indices of order k when β runs over all multi-indices of order $i = k - j$ and j runs from 0 to k, we obtain

$$\sum_{|\alpha|=k} \frac{k! \boldsymbol{x}^\alpha}{\alpha!}.$$

□

Analogously we are lead to the product rule for higher-order partial derivatives

$$\partial^\alpha (fg) = \sum_{\beta+\gamma=\alpha} \frac{\alpha!}{\beta!\gamma!} (\partial^\beta f)(\partial^\gamma g).$$

5.1.3 Deriving the multi-variable case

We shall now finally turn to Taylor's theorem for functions of several variables. For simplicity scalar-valued functions are considered first and vector-valued case is presented right after.

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^k on a convex open set S. We derive a Taylor expansion for $f(\boldsymbol{x})$ about a point $\boldsymbol{a} \in S$ by looking at the restriction of f to the line joining $\boldsymbol{a}$ and $\boldsymbol{x} \in S$ That is, we set $\boldsymbol{h} = \boldsymbol{x} - \boldsymbol{a}$ and

$$g(t) = f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})) = f(\boldsymbol{a} + t\boldsymbol{h}).$$

By the chain rule,

$$g'(t) = \boldsymbol{h} \cdot \nabla f(\boldsymbol{a} + t\boldsymbol{h}),$$

thus

$$g^{(j)}(t) = (\boldsymbol{h} \cdot \nabla)^j f(\boldsymbol{a} + t\boldsymbol{h}),$$

where the expression on the right denotes the result of applying the directional derivate

$$\boldsymbol{h} \cdot \nabla = h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \tag{5.2}$$

j times to f. The Taylor formula for g around $t = 0$ and evaluated at $h = 1$ reads

$$g(1) = \sum_{j=0}^{k} \frac{g^j(0)}{j!} 1^j + \text{remainder},$$

which yields

$$f(\boldsymbol{a} + \boldsymbol{h}) = \sum_{j=0}^{k} \frac{(\boldsymbol{h} \cdot \nabla)^j f(\boldsymbol{a})}{j!} + R_{a,k}(\boldsymbol{h}), \tag{5.3}$$

where $R_{a,k}(\boldsymbol{h})$ denotes the remainder term.

It is usually preferable to rewrite (5.3) including the accompanying formula for the remainder so that the partial derivatives of f appear more explicitly. To do this, we apply the multinomial theorem to the expression (5.2) to get

$$(\boldsymbol{h} \cdot \nabla)^j = \sum_{|\alpha|=j} \frac{j!}{\alpha!} \boldsymbol{h}^\alpha \partial^\alpha. \tag{5.4}$$

When this is substituted back into (5.3) we obtain Taylor's theorem in several variables [Hol, p. 1-2].

5.1.4 Taylor's theorem for multi-variable scalar-valued functions

Theorem 5.1.2. *(Taylor's multi-variable theorem for **scalar**-valued functions). Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^{k+1} on an open convex set $S \in \mathbb{R}^n$. If $\boldsymbol{a} \in S$ and $\boldsymbol{a} + \boldsymbol{h} \in S$, then*

$$f(\boldsymbol{a} + \boldsymbol{h}) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(\boldsymbol{a})}{\alpha!} \boldsymbol{h}^\alpha + R_{\boldsymbol{a},k}(\boldsymbol{h}), \tag{5.5}$$

where the remainder is given in Lagrange's form by

$$R_{\boldsymbol{a},k}(\boldsymbol{h}) = \sum_{|\alpha|=k+1} \partial^\alpha f(\boldsymbol{a} + c\boldsymbol{h}) \frac{\boldsymbol{h}^\alpha}{\alpha!} \text{ for some } c \in (0,1). \tag{5.6}$$

and in integral form by

$$R_{\boldsymbol{a},k}(\boldsymbol{h}) = (k+1) \sum_{|\alpha|=k+1} \frac{\boldsymbol{h}^\alpha}{\alpha!} \int_0^1 (1-t)^k \partial^\alpha f(\boldsymbol{a} + t\boldsymbol{h}) dt. \tag{5.7}$$

A specified corollary to estimate the remainder follows from (5.6) or (5.7).

Corollary 5.1.2.1. *If f is of class C^{k+1} on $S \in \mathbb{R}^n$, $|\partial^\alpha f(\boldsymbol{x})| \leq M \geq 0$ for $\boldsymbol{x} \in S$ and $|\alpha| = k+1$, then*

$$|R_{\boldsymbol{a},k}(\boldsymbol{h})| \leq \frac{M}{(k+1)!} \parallel \boldsymbol{h} \parallel^{k+1},$$

where

$$\parallel \boldsymbol{h} \parallel = |h_1| + |h_2| + \cdots + |h_n|.$$

5.1.5 Taylor's theorem for multi-variable vector-valued functions

A scalar-valued function f can be expressed as a combination of its components

$$f(\boldsymbol{a}+\boldsymbol{h}) = \sum_{j=0}^{k}\sum_{i=1}^{n}\frac{(\boldsymbol{h}\cdot\nabla)^j f_i(\boldsymbol{a})}{j} + R_{(a,k)_i}(\boldsymbol{h}). \tag{5.8}$$

Next we take the sum over the dot product of each unit vector e_i with each component of (5.8) and land with vector-valued case of Taylor's multi-variable theorem.

Theorem 5.1.3. *(Taylor's multi-variable theorem for **vector**-valued functions). Suppose $f : \mathbb{R}^m \to \mathbb{R}^n$ is of class C^{k+1} on an open convex set $S \in \mathbb{R}^n$. Now as $\boldsymbol{a} \in S$ and $\boldsymbol{a}+\boldsymbol{h} \in S$ vector-valued function $\boldsymbol{f}$ can be expressed as*

$$\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h}) = \sum_{i=1}^{n} f_i(\boldsymbol{a}+\boldsymbol{h})\boldsymbol{e}_i = \sum_{i=1}^{n}\left(\sum_{|\alpha|\le k}\frac{\partial^\alpha f_i(\boldsymbol{a})}{\alpha!}\boldsymbol{h}^\alpha + R_{\boldsymbol{a},k,i}(\boldsymbol{h}))\right)\boldsymbol{e}_i, \tag{5.9}$$

where $\boldsymbol{e}_i$ is a unit vector pointing towards dimension i.

Corollary 5.1.3.1. *If $\boldsymbol{f}$ is of class C^{k+1} on $S \in \mathbb{R}^n$ and $|\partial^\alpha \boldsymbol{f}(\boldsymbol{x})| \le M \ge 0$ for $\boldsymbol{x} \in S$ and $|\alpha| = k+1$, $\forall\alpha$, then*

$$|R_{\boldsymbol{a},k}(\boldsymbol{h})| \le \frac{M}{(k+1)!} \parallel \boldsymbol{h} \parallel^{k+1},$$

where

$$\parallel \boldsymbol{h} \parallel = |h_1| + |h_2| + \cdots + |h_n|.$$

For additional reading about Taylor's theorem and differentiation of vector valued functions Walter Rudin's classic Principles of Mathematical Analysis [Rud76] is highly recommended.

5.2 Jacobian

For a one-dimensional function of one variable we have

$$f'(x_0) \approx \frac{f(x_0+\epsilon)-f(x_0)}{\epsilon} \quad \text{for } \epsilon \ll 1,$$

with which the derivative at $x = x_0$ can be approximated. The higher-dimensional generalization of this is the Jacobian matrix, which describes the local linear behavior of a higher-dimensional function. Jacobian can be seen as a derivative of a coordinate transformation. In a matrix form Jacobian for function $f : \mathbb{R}^n \to \mathbb{R}^m$ is defined as

$$J_f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

and for function $f : \mathbb{R}^3 \to \mathbb{R}^3$ as

$$J_f = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \end{bmatrix}$$

[Har03, p. 6-8].

5.3 Linearization

In this section we present the concept of linearization in $\mathbb{R}^3$. Consider an autonomous vector field $f : \mathbb{R}^3 \to \mathbb{R}^3$

$$\boldsymbol{f}(\boldsymbol{y}) = \frac{d\boldsymbol{y}}{dt}. \tag{5.10}$$

In regards of vector fields we can speak of dynamical systems in a similar fashion as in Chapter 4. We want to understand the nature of solutions near $\boldsymbol{y}(0)$, which is usually considered as the initial point from which the observation begins. The point of interest is $\boldsymbol{y}(t)$. Let $\boldsymbol{h}(t)$ be the vector indicating the difference between $\boldsymbol{y}(0)$ and $\boldsymbol{y}(t)$. See Figure 5.1.

Thus we have

$$\boldsymbol{y}(t) = \boldsymbol{y}(0) + \boldsymbol{h}(t) \tag{5.11}$$

Now by substituting (5.10) to (5.11) we have

$$\frac{d}{dt}\boldsymbol{y}(t) \quad = \quad \boldsymbol{f}\left(\boldsymbol{y}(0) + \boldsymbol{h}(t)\right) \tag{5.12}$$

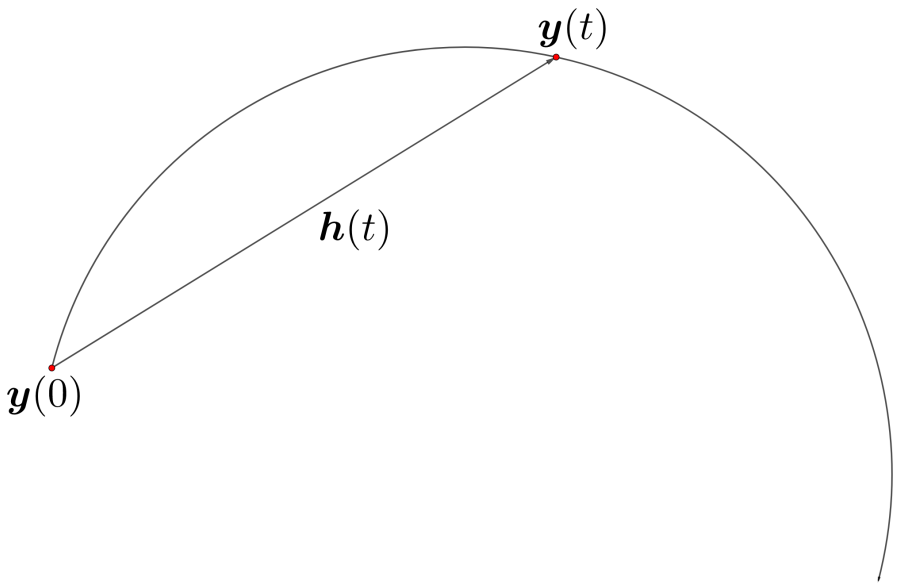


Figure 5.1: Linearization. The shorter the $\boldsymbol{h}(t)$ the better the approximation for $\boldsymbol{y}(t)$.

Next we write a step-by-step Taylor expansion for (5.12) following closely Theorem 5.1.3.

(5.13)
$$\begin{aligned}
\boldsymbol{f}(\boldsymbol{y}(0)+\boldsymbol{h}(t)) &= \sum_{|\alpha|\leq 1} \frac{\partial^{\alpha}\boldsymbol{f}(\boldsymbol{y}(0))}{\alpha!}\boldsymbol{h}^{\alpha}(t) + R_{\boldsymbol{y}(0),1}(\boldsymbol{h}(t)) \\
&= \frac{\partial^{(0,0,0)}\boldsymbol{f}(\boldsymbol{y}(0))}{0!0!0!}\boldsymbol{h}^{(0,0,0)}(t) + \frac{\partial^{(1,0,0)}\boldsymbol{f}(\boldsymbol{y}(0))}{1!0!0!}\boldsymbol{h}^{(1,0,0)}(t) \\
&\quad + \frac{\partial^{(0,1,0)}\boldsymbol{f}(\boldsymbol{y}(0))}{0!1!0!}\boldsymbol{h}^{(0,1,0)}(t) \quad + \frac{\partial^{(0,0,1)}\boldsymbol{f}(\boldsymbol{y}(0))}{0!0!1!}\boldsymbol{h}^{(0,0,1)}(t) \\
&\quad + R_{\boldsymbol{y}(0),1}(\boldsymbol{h}(t)) \\
&= \boldsymbol{f}(\boldsymbol{y}(0)) + \frac{\partial}{dx}\boldsymbol{f}(\boldsymbol{y}(0))\boldsymbol{h}_1(t) + \frac{\partial}{dy}\boldsymbol{f}(\boldsymbol{y}(0))\boldsymbol{h}_2(t) + \frac{\partial}{dz}\boldsymbol{f}(\boldsymbol{y}(0))\boldsymbol{h}_3(t) \\
&\quad + R_{\boldsymbol{y}(0),1}(\boldsymbol{h}(t))).
\end{aligned}$$

The above can be concisely written as

$$\boldsymbol{f}(\boldsymbol{y}(t)) = \boldsymbol{f}(\boldsymbol{y}(0)) + J_f(\boldsymbol{y}(t))\boldsymbol{h}(t) + R_{\boldsymbol{y}(0),1}(\boldsymbol{h}(t))), \tag{5.14}$$

where J_f is a 3×3-Jacobian matrix [Wig03, p. 10-11].

Chapter 6

Lorenz System

6.1 Introducing Lorenz system

Lorenz described in his remarkable 1963 paper a set of three-dimensional ordinary differential three parameter equations. Lorenz's initial interest was to model some of the unpredictable behavior normally associated with weather. Although Lorenz system is known as *the system* for modeling weather it is simplified to an extent that it can not be used for studying extended weather systems. Rather, Lorenz system apply to a single convection cell illustrated in Figure 6.1.

The original derivation of the equations can be described briefly as follows. A two-dimensional fluid cell is cooled from top and heated from bottom and the arising convective motion is modeled by a partial differential equation. It is good to bear in mind that Lorenz equations can be studied solely as a mathematically formulated system ignoring the context of weather modeling. Regardless of the context the famous Lorenz equations in their full glory are written as

$$\begin{aligned} \frac{\partial x}{\partial t} &= a(y - x) \\ \frac{\partial y}{\partial t} &= x(c - z) - y \\ \frac{\partial z}{\partial t} &= xy - bz. \end{aligned} \tag{6.1}$$

In the context of the original weather model the variable x measures the rate of convective overturning, the variable y measures the horizontal temperature and the variable z measures the vertical temperature variation. The three parameters a, b and c are respectively

proportional to Prandtl number, the Rayleigh number, and some physical proportions of the region under consideration [Spa12, p. 1-2]. In our study we focus only on parameter values $a = 10$, $b = 28$ and $c = 8/3$ which yield the following equations

$$\begin{aligned} \frac{\partial x}{\partial t} &= 10(y - x) \\ \frac{\partial y}{\partial t} &= x(\frac{8}{3} - z) - y \\ \frac{\partial z}{\partial t} &= xy - 28z. \end{aligned} \tag{6.2}$$

More specifically described, Lorenz equations model convection in which fluid is constrained to flow in a toroidal loop that is heated at the bottom and cooled at the top. The hot fluid is lighter than the cold, so in the absence of initial motion we have an unstable situation. The hot, buoyant fluid at the bottom of the loop would like to rise and displace the cold, dense fluid at the top. The slightest force will initiate the fluid to rotate in one direction or the other [Kau11, p. 146-147].

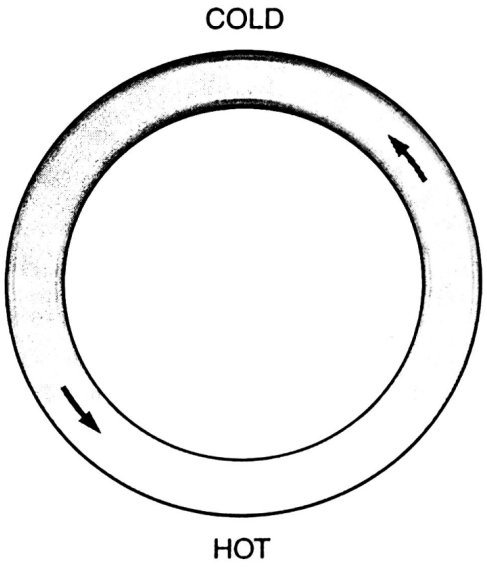


Figure 6.1: Convection Loop [PJS04, p. 146].

6.2 A Vector Field

6.2.1 Definition

In contexts such as vector calculus, physics, and dynamical systems, a vector field is concept where in a subset of a space a vector is assigned to each point. See figure 6.2. A vector field can be visualized as a collection of arrows with a given magnitude and direction. Vector fields are used to model different systems in nature. Such systems range from fluid moving throughout space to different magnetic or gravitational fields. It's important to notice that the magnitude and the direction of a vector in a vector field can change with respect to its position.

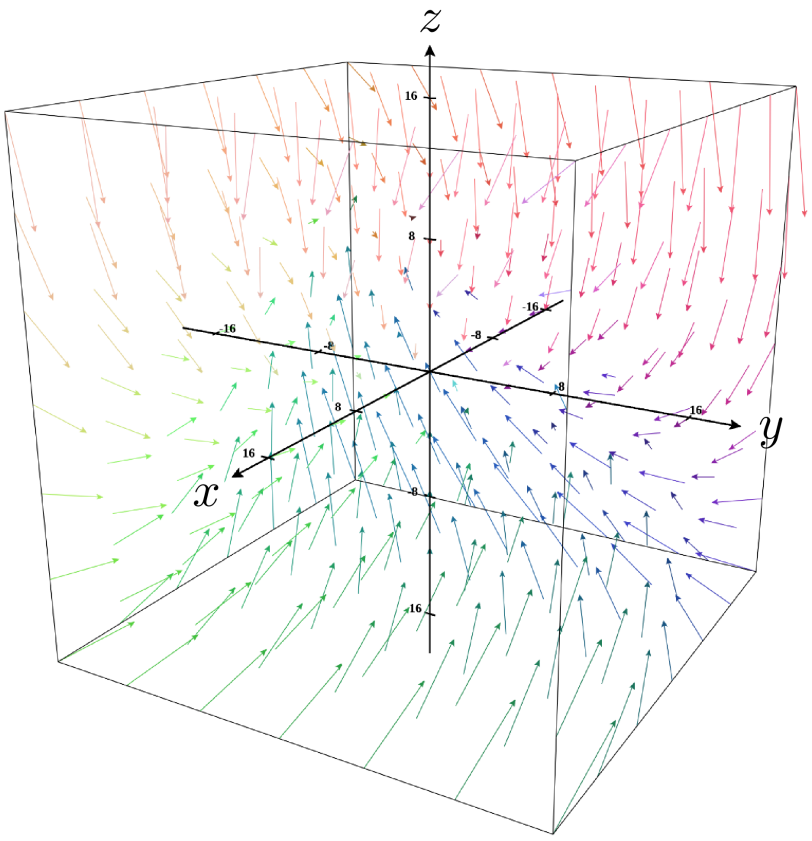


Figure 6.2: Vector Field Presentation of Lorenz System. Note how every point is assigned with a vector. Drawn with 3D plotter [3DP].

In the context of coordinate systems, a vector field on n-dimensional Euclidean space can

be represented as a vector-valued function that partners an n-tuple of real numbers to each point of the space. Vector fields are often discussed on open subsets of Euclidean space but they could as well be operating on other subsets such as surfaces [Joh08, p. 2-3].

6.2.2 Lorenz System as a Vector Field

Let us consider $\boldsymbol{y}(t) = (x, y, z) \in \mathbb{R}^3$ as the solution to (6.1). Next we define that

$$\frac{d}{dt}\boldsymbol{y}(t) = \boldsymbol{f}\left(\boldsymbol{y}(t)\right). \tag{6.3}$$

In Figure 6.3 we see that besides constituting of components $\frac{\partial x}{\partial t}$, $\frac{\partial y}{\partial t}$ and $\frac{\partial z}{\partial t}$ the following vector field

$$\boldsymbol{f}(\boldsymbol{y}) = \begin{pmatrix} 10(y-x) \\ x(\frac{8}{3}-z)-y \\ xy-28z \end{pmatrix} \tag{6.4}$$

assigns a vector to each point (x, y, z). One can picture the situation as *if within a system one knows its location one knows its direction.*

When regarding dynamical systems a vector field should be seen as the fixed system of rules that defines future behavior of a particular state. In other words $\boldsymbol{f}(\boldsymbol{y})$ defines the trajectory of any initial conditions. Accordingly if we assume that $\boldsymbol{y}(0)$ is known we can treat $\boldsymbol{y}(t)$ as the vector field's flow, a concept introduced in Chapter 4.

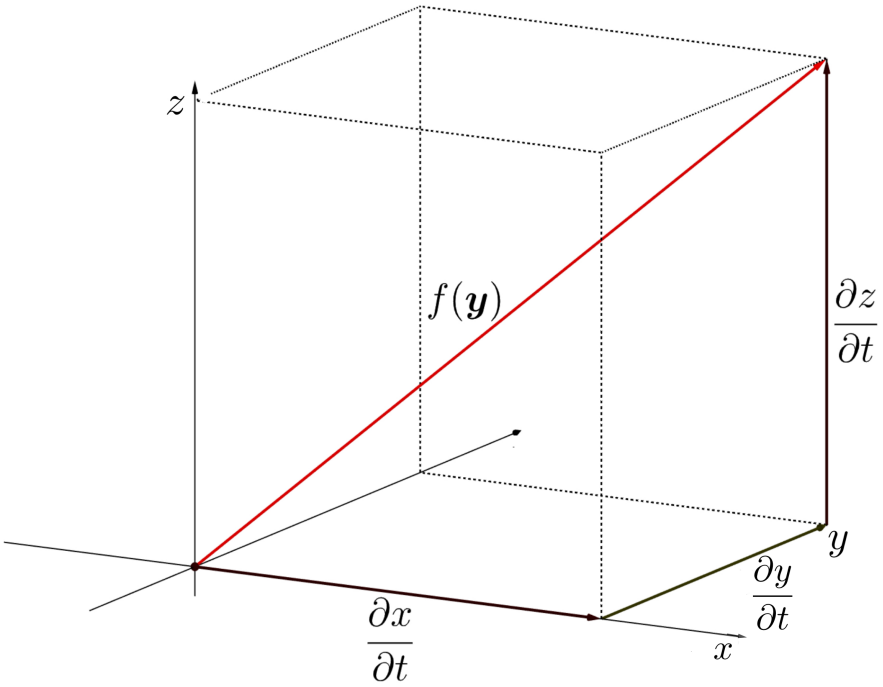


Figure 6.3: Each vector in a vector field constitutes of its components. In $\mathbb{R}^3$ one can imagine a vector $\boldsymbol{f}(\boldsymbol{y})$ attached to every point $\boldsymbol{y} = (x, y, z)$.

6.3 Geometrical Representation of the Solution Path

Figure 6.4 is a graphical representation of the flow $\boldsymbol{y}(t)$ over time starting from an initial condition of $\boldsymbol{y}(0) = (0, 1, 0)$, which can be identified as the point outside of the two attractor surfaces on which the plot seems to circulate relatively periodically. The layman's way to read trajectory's path is to picture a situation where an object is thrown to a system which looks like Figure 6.2. As it is by definition obeying the system it will follow a route presented in Figure 6.4. This visual metaphor obviously works only in dimensions where it is possible to visualize the trajectory.

The butterfly looking attractor the trajectory seems to lay on, as seen in Figure 6.4, peculiarly pull the trajectory to close distance no matter how far the initial condition $\boldsymbol{y}(0)$ is set. This and other less quantifiable properties of the Lorenz system are discussed more in Chapter 8.

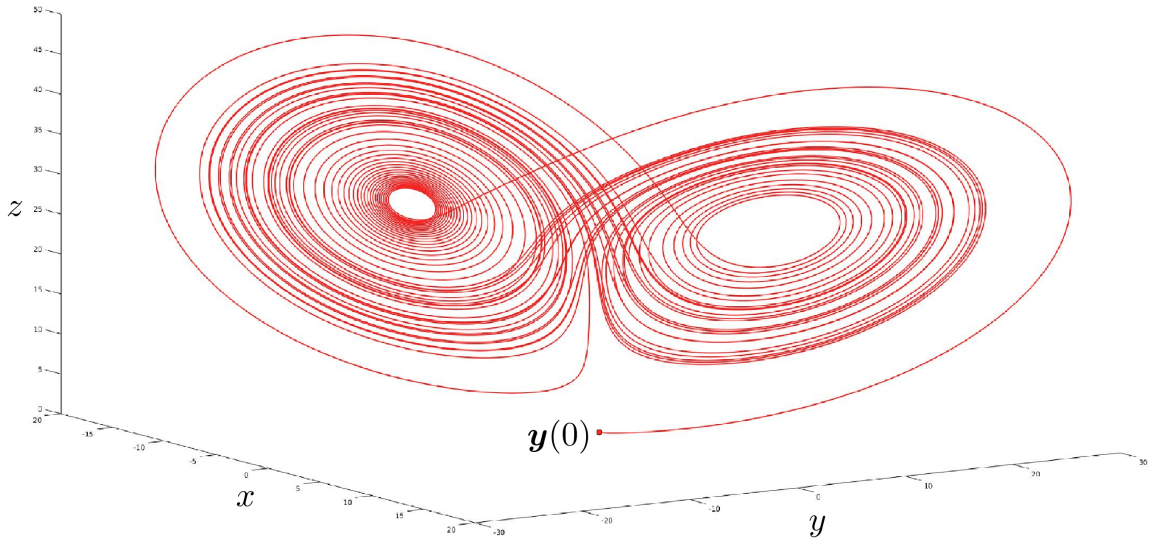


Figure 6.4: Lorenz attractor. A numerically computed solution to the Lorenz equations when $a = 10$, $b = 28.0$ and $c = 8/3$. Computed with MATLAB, $dt = 0.001$, timespan $0.001 - 50$ and initial condition $\boldsymbol{y}(0) = (0, 1, 0)$.

Chapter 7

Lorenz System's Sensitivity to Initial Conditions

In this chapter we examine Lorenz system's sensitive dependence on initial conditions. When a system exhibits sensitive dependence on initial conditions two trajectories starting very close together rapidly diverge from each other, and thereafter have totally different futures paths. The observational implication is that in a sensitively dependent system, where small uncertainties are amplified enormously fast, long-term prediction becomes impossible. In the following we show that such behavior emerge in Lorenz system. To do this we first define a general deviance between two initially nearby trajectories and then we study the evolution of this deviance [Str94, p. 320-322].

7.1 Time Evolution of Deviance

We observe two trajectories $\boldsymbol{y}^{(1)}(t)$ and $\boldsymbol{y}^{(2)}(t)$, $\boldsymbol{y}^{(1)}, \boldsymbol{y}^{(2)} \in \mathbb{R}^3$. See Figure 7.1 for illustration. If we define $\boldsymbol{\delta}(t)$ as the deviance between $\boldsymbol{y}^{(1)}(t)$ and $\boldsymbol{y}^{(2)}(t)$ we have by definition

$$\boldsymbol{y}^{(2)}(t) = \boldsymbol{y}^{(1)}(t) + \boldsymbol{\delta}(t), \tag{7.1}$$

and especially when $t = 0$

$$\boldsymbol{y}^{(2)}(0) = \boldsymbol{y}^{(1)}(0) + \boldsymbol{\delta}(0). \tag{7.2}$$

Now we approximate $\boldsymbol{y}^{(2)}$ in the same fashion as in (5.13)

$$\begin{aligned} \boldsymbol{f}(\boldsymbol{y}^{(2)}(t)) &= \boldsymbol{f}(\boldsymbol{y}^{(1)}(t) + \boldsymbol{\delta}(t)) \\ &= \boldsymbol{f}(\boldsymbol{y}^{(1)}(t)) + J_f(\boldsymbol{y}^{(1)}(t))\boldsymbol{\delta}(t) + R_{\boldsymbol{y}^{(1)}(t),1}(\boldsymbol{\delta}(t)). \end{aligned} \tag{7.3}$$

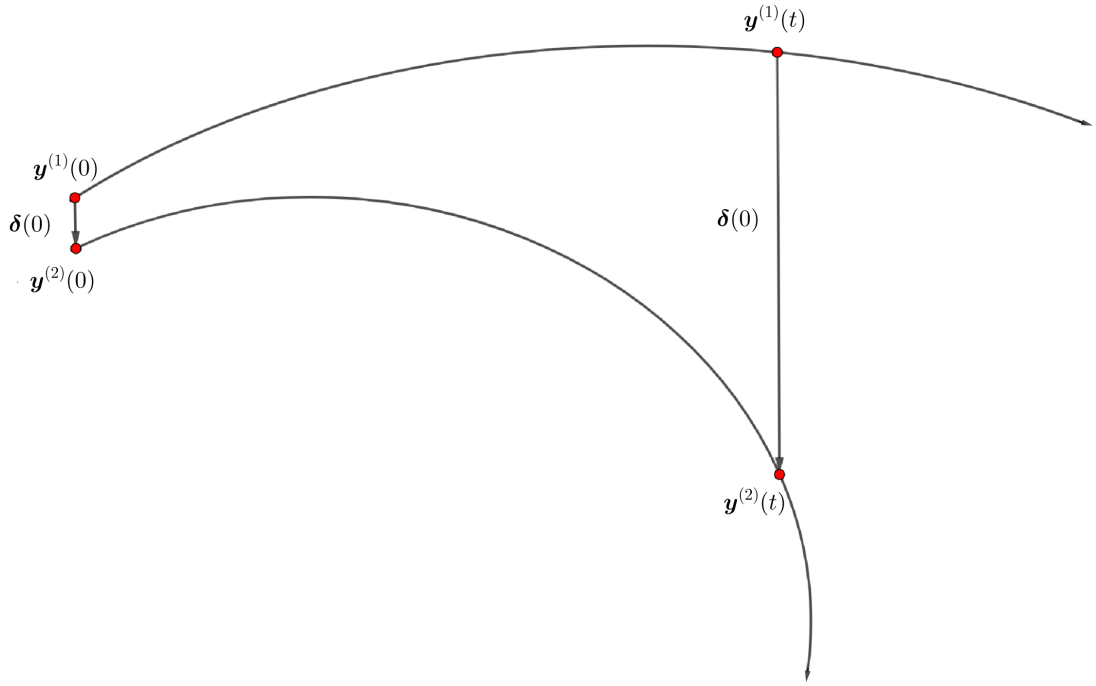


Figure 7.1: Divergence of two initially nearby trajectories.

and now by combining (7.1) with (7.3) we have

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{\delta}(t)) &= \boldsymbol{f}(\boldsymbol{y}^{(2)}(t) - \boldsymbol{y}^{(1)}(t)) \\ &= \boldsymbol{f}(\boldsymbol{y}^{(2)}(t)) - \boldsymbol{f}(\boldsymbol{y}^{(1)}(t)) \\ &= \boldsymbol{f}(\boldsymbol{y}^{(1)}(t)) + J_f(\boldsymbol{y}^{(1)}(t))\boldsymbol{\delta}(t) + R_{\boldsymbol{y}^{(1)}(t),1}(\boldsymbol{\delta}(t)) - \boldsymbol{f}(\boldsymbol{y}^{(1)}(t)) \\ &= J_f(\boldsymbol{y}^{(1)}(t))\boldsymbol{\delta}(t) + R_{\boldsymbol{y}^{(1)}(t),1}(\boldsymbol{\delta}(t)).\end{aligned} \tag{7.4}$$

Now we have a general equation for a time evolution of a vector defining the deviation of two trajectories in $\mathbb{R}^3$.

$$\frac{d}{dt}\boldsymbol{\delta}(t) = J_f(\boldsymbol{y}^{(1)}(t))\boldsymbol{\delta}(t) + R_{\boldsymbol{y}^{(1)}(t),1}(\boldsymbol{\delta}(t)). \tag{7.5}$$

7.1.1 Remainder Estimation

Now we use the Corollary 5.1.3.1 to estimate the remainder term $R_{\boldsymbol{y}^{(1)}(t),1}(\boldsymbol{\delta}(t))$. First we note that $\boldsymbol{f}$, defined in (6.3), is of class C^{k+1}. Thus we have $\boldsymbol{y} \in \mathbb{R}^3$ and $|\alpha| = 2$ as $k = 1$. Next we start estimating M. We need to find M so that $|\partial^\alpha \boldsymbol{f}(\boldsymbol{y})| \leq M$ holds. Hence we find the maximum value for $|\partial^\alpha \boldsymbol{f}(\boldsymbol{y})|$.

By definition we have

$$\partial^\alpha \boldsymbol{f} = \frac{\partial^{|\alpha|} \boldsymbol{f}}{\partial y_1^{\alpha_1} \partial y_2^{\alpha_2} \cdots \partial y_n^{\alpha_n}},$$

and possible options for α are $(1,1,0), (1,0,1), (0,1,1), (0,0,2), (0,2,0), (2,0,0)$.

With all α options other than $\alpha = (1,1,0)$ we have $\partial^\alpha f = 0$. For illustration we compute the value for $|\partial^\alpha \boldsymbol{f}(\boldsymbol{y})|$ with $\alpha = (1,1,0)$.

$$\begin{aligned}
\left| \frac{\partial^{|\alpha|} \boldsymbol{f}}{\partial y_1^{\alpha_1} \partial y_2^{\alpha_2} \cdots \partial y_n^{\alpha_n}} \right| &= \left| \frac{\partial^2 \boldsymbol{f}}{\partial x^1 \partial y^1 \partial z^0} \right| \\
&= \left| \frac{\partial^2 \boldsymbol{f}}{\partial x \partial y} \right| \\
&= |(0,0,1)| && (\boldsymbol{f} \text{ defined in (6.4)}) \\
&= \sqrt{1^2} && \left(|\boldsymbol{f}|_2 := \sqrt{\sum_{i=1}^{n} |f_i^2|} \right) \\
&= 1.
\end{aligned}$$

We can set $M = 2$ and focus on relatively short initial deviation vector values $\boldsymbol{h} = \boldsymbol{\delta}(0) = (\frac{1}{10^6}, \frac{1}{10^6}, \frac{1}{10^6})$. Now the magnitude of the remainder can be estimated.

$$\begin{aligned}
|R_{\boldsymbol{a},k}(\boldsymbol{\delta}(0))| &\leq \frac{M}{(k+1)!} \parallel \boldsymbol{\delta}(0) \parallel^{k+1} \\
&= \frac{2}{2!} \parallel \boldsymbol{\delta}(0) \parallel^2 && (\text{as } k = 1 \text{ and } M = 2) \\
&= \frac{2}{2} \left(\left| \frac{1}{10^6} \right| + \left| \frac{1}{10^6} \right| + \left| \frac{1}{10^6} \right| \right)^2 && (\text{Corollary 5.1.3.1 defines } \parallel \boldsymbol{\delta}(0) \parallel \text{ as } l_1 \text{ norm}) \\
&= 9 \times 10^{-12}.
\end{aligned}$$

Hereby we have a remainder $R_{\boldsymbol{a},k}(\boldsymbol{\delta}(0))$ with magnitude of less than 9×10^{-12} giving margin for $\boldsymbol{f}$'s Taylor series approximation.

7.1.2 Deviation In Tangent Space

As the error margin is so small we ignore the higher order terms by identifying a deviation vector $\boldsymbol{\delta}$ with an element $\boldsymbol{\xi} \in \mathbb{R}^3$ in the tangent space. Denoting $\boldsymbol{\xi}(t) = \boldsymbol{\delta}(t)$ we are left to solve

$$\frac{d}{dt}\boldsymbol{\xi}(t) \quad = \quad J_f(\boldsymbol{y}^{(1)}(0))\boldsymbol{\xi}(t), \tag{7.6}$$

which is a first-order differential equation. To clarify this we write

$$\boldsymbol{\xi}'(t) \quad = \quad J_f(\boldsymbol{y}^{(1)}(0))\boldsymbol{\xi}(t). \tag{7.7}$$

As explained in Appendix A.2 the solution to (7.7) is of form

$$\boldsymbol{\xi}(t) \quad = \quad e^{J_f(\boldsymbol{y}^{(1)}(0))t}\boldsymbol{\xi}(0). \tag{7.8}$$

Because of the Jacobian matrix $J_f\left(\boldsymbol{y}^{(1)}(0)\right)$ in the exponent the values of (7.8) are less straight-forward to compute. This is why the following section presents a simpler way of approximating the deviance $\boldsymbol{\xi}(t)$ [Har03, p. 8].

7.2 Deviance Vector Length Estimation

We study the length of $\boldsymbol{\xi}(t)$ projected to its Jacobian transpose matrix's eigenvector $\boldsymbol{v}$. To do this we set a definition regarding the Jacobian matrix and its eigenvalues and prove a specific proposition which will help us to estimate deviance vector's length.

Definition 7.2.1. Let $\lambda \in \mathbb{C}$ be the eigenvalue and $\boldsymbol{v} \neq 0$ the eigenvector of transpose-matrix A^T. We denote $J_f(\boldsymbol{y_0}) = A$ and set the following postulates

Postulate 1. $\frac{d}{dt}\boldsymbol{\xi}(t) \quad = \quad A\boldsymbol{\xi}(t)$

Postulate 2. $A^T\boldsymbol{v} = \lambda\boldsymbol{v} \quad$ and $\quad |\boldsymbol{v}| = 1.$

We are interested in the behavior of $\boldsymbol{\xi}(t)$ with relation to $\boldsymbol{v}$. Thus we denote

$$g(t) := \boldsymbol{v} \cdot \boldsymbol{\xi}(t) \tag{7.9}$$

and focus on function $g(t)$. Next is shown that the time evolution $\frac{d}{dt}g(t)$ can be approximated with a scalar-multiplier λ. In other words the following proposition is shown to be true.

Proposition 1. *If Definition 7.2.1 including its postulates I and II is true, then* $\frac{d}{dt}g(t) = \lambda g(t)$ *is also true.*

Proof.

$$\begin{aligned}
\frac{d}{dt}g(t) &= \frac{d}{dt}\left(\boldsymbol{v} \cdot \boldsymbol{\xi}(t)\right) && (7.9)\\
&= \frac{d}{dt}\left(\sum_{i=1}^{3} v_i \xi_i(t)\right) && (\text{as } \xi_i \in \mathbb{R})\\
&= \sum_{i=1}^{3} v_i \frac{d}{dt} \sum_{j=1}^{3} \xi_j \\
&= \sum_{i=1}^{3} v_i \sum_{j=1}^{3} A_{ij}\xi_j(t) && (P.I)\\
&= \sum_{j=1}^{3} \xi_j \sum_{i=1}^{3} v_i (A^T)_{ji} \\
&= \sum_{j=1}^{3} \xi_j (A^T \boldsymbol{v})_j \\
&= \boldsymbol{\xi}(t) \cdot A^T \boldsymbol{v} \\
&= \boldsymbol{\xi}(t) \cdot \lambda \boldsymbol{v} && (P.II)\\
&= \lambda \boldsymbol{v} \cdot \boldsymbol{\xi}(t) \\
&= \lambda g(t).
\end{aligned}$$

□

Hence we have function

$$g(t) = \boldsymbol{v} \cdot \boldsymbol{\xi}(t) \tag{7.10}$$

satisfying

$$g'(t) = \lambda g(t). \tag{7.11}$$

(7.11) is similarly to (7.7) a first-order differential equation. And as $g(0)$ is known, the equation is easily solved as an *initial value problem*

$$\begin{aligned} g(t) &= e^{\lambda t} g(0) \\ &= e^{\lambda t} \boldsymbol{v} \cdot \boldsymbol{\xi}(0). \qquad \text{(see 7.10)} \end{aligned}$$

In summary, we have

$$\boldsymbol{v} \cdot \boldsymbol{\xi}(t) = e^{\lambda t} \boldsymbol{v} \cdot \boldsymbol{\xi}(0). \tag{7.12}$$

Since $|\boldsymbol{v}| = 1$ the Cauchy-Schwarz inequality explained in Appendix A.3 shows us that $|\boldsymbol{\xi}(t) \cdot \boldsymbol{v}| \leq |\boldsymbol{\xi}(t)|$. Hence if $c_0 := |\boldsymbol{v} \cdot \boldsymbol{\xi}(0)|$ and $\lambda \in \mathbb{C}$, we can write

$$\begin{aligned} |\boldsymbol{\xi}(t)| &\geq |\boldsymbol{v} \cdot \boldsymbol{\xi}(t)| \\ &= |e^{\lambda t} \boldsymbol{v} \cdot \boldsymbol{\xi}(0)| \\ &= c_0 e^{Re(\lambda)t}, \end{aligned}$$

as the imaginary part of $e^{\lambda t}$ can be ignored. See Appendix A.4 for further interest. Now because $c_0 := |\boldsymbol{v} \cdot \boldsymbol{\xi}(0)|$ is a constant determined by initial conditions we can write

$$|\boldsymbol{\xi}(t)| \geq c_0 e^{Re(\lambda)t}. \tag{7.13}$$

7.3 Computing Jacobian's Eigenvalues at an Arbitrary Point

Inequality (7.13) is saying that if we can find a positive eigenvalue λ for matrix A^T we can estimate the deviance $\boldsymbol{\xi}(t)$ to grow exponentially after a certain point in time t whenever $c_0 \neq 0$. See Section 7.2 for c_0. In other words the estimation works for most initial data. This is why we try to find eigenvalues λ with positive real part for matrix A^T. By denoting $A^T = J_f(\boldsymbol{y}(0))^T$ we can formulate the following proposition.

Proposition 2. $\exists \boldsymbol{y}(0) \in \mathbb{R}^3$ *and* $\exists \lambda > 0$ *that simultaneously satisfy* $A^T \boldsymbol{y}(0) = \lambda \boldsymbol{y}(0)$.

Proof. We set the initial point to $\boldsymbol{y}(0) = (0.001, 0.001, 0.001)$ yielding

$$A^T = \begin{bmatrix} -10 & 28 - 10^{-3} & 10^{-3} \\ 10 & -1 & 10^{-3} \\ 0 & -10^{-3} & -8/3 \end{bmatrix}. \tag{7.14}$$

MATLAB computed eigenvalues for A^T at $\boldsymbol{y}(0)$ are

$$\begin{aligned} \lambda_1 &= -22.82... \\ \lambda_2 &= 11.82... \\ \lambda_3 &= -2.66... \end{aligned}$$

□

The above shows how a positive eigenvalue λ_2 can be found. This is obviously not too informative as it shows that a positive eigenvalue can be found in just one point but does not give information about system's initial value dependency on a wider scale. In the next section we take a look if we can broaden our scope and see where in $\mathbb{R}^3$ we can find positive eigenvalues.

7.4 Is Lorenz System Sensitive to Initial Conditions Everywhere?

Where in $\mathbb{R}^3$ the Lorenz system shows sensitivity to initial conditions? We search for regions in $\mathbb{R}^3$ where there is an eigenvalue with a positive real part.

Figure 7.2 demonstrates the distribution of maximum eigenvalues of J_f^T in a constrained but relatively wide xy-plane. Yellow color marks for positive eigenvalues and blue color for negative eigenvalues. Essentially Figure 7.2 shows us that Lorenz system displays positive eigenvalues on a significant region of $\mathbb{R}^3$. Within this region the Lorenz system shows chaotic behavior to at least one direction of $\mathbb{R}^3$. Thus we can ascertain how the Lorenz system truly is sensitive to initial conditions on certain significant regions on $\mathbb{R}^3$ but not everywhere.

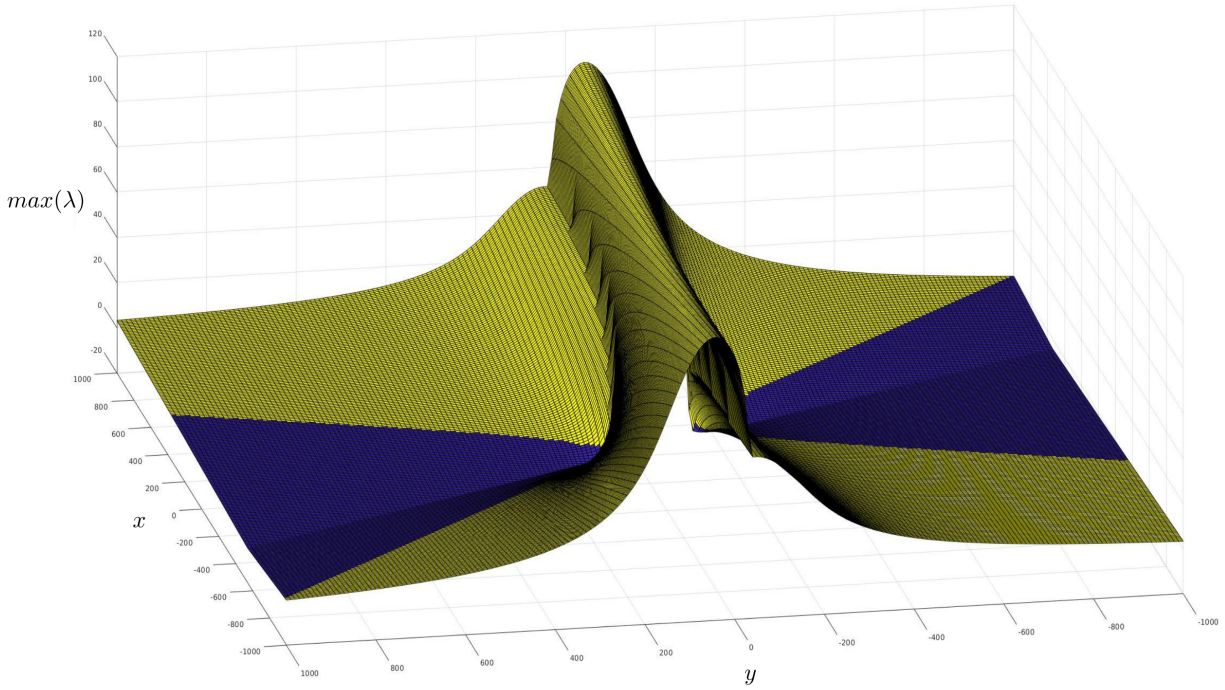


Figure 7.2: Maximum eigenvalue of Jacobian transpose matrix J_f^T plotted against xy-plane. Positive values are marked with yellow color and negative values with blue color.

7.5 Prediction Breaks Down Figuratively

In this section the Lorenz system's unpredictability is figuratively demonstrated. In the following we follow the development of three Lorenz trajectories over a short period of time.

Figures 7.3 and 7.4 demonstrate the development of three Lorenz trajectories with different initial conditions. The blue trajectory starts from an initial conditions $\boldsymbol{y}_B(0) = (0, 1, 0)$, the red trajectory starts from $\boldsymbol{y}_R(0) = (0, 0.9, 0)$ and the green from $\boldsymbol{y}_G(0) = (0, 0.99, 0)$. Note that green's and blue's initial points are closer together than blue's and red's. In the plots of Figure 7.3 the initial points can be observed loosely in the lower end of the plot of each graph. As the initial conditions are relatively close together the initial points are marked with a point BRG in Figure 7.4 referring to blue, red and green. Note that all

three initial points $\boldsymbol{y}_B(0), \boldsymbol{y}_R(0)$ and $\boldsymbol{y}_G(0)$ are demonstrated as one point BRG as they are so close to each other.

7.5.1 Fixed Points

Before describing the behavior of the three trajectories the term *fixed point* needs to be defined. Function's steady solutions, or fixed points, occur when

$$\frac{\partial x}{\partial t} = \frac{\partial y}{\partial t} = \frac{\partial z}{\partial t} = 0 \tag{7.15}$$

For the Lorenz system one obvious fixed point is when

$$x = y = z = 0.$$

The other two steady solutions are

$$x = y = \sqrt{c-1} \text{ and } z = c - 1$$

and

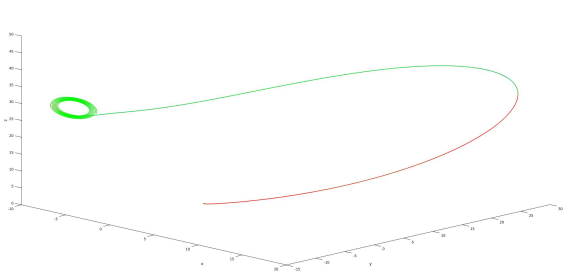
$$x = y = -\sqrt{c-1} \text{ and } z = c - 1.$$

Geometrically fixed points of the Lorenz system can be observed for example in Figure 7.4 as the origin, which has no obvious role in the graph, and the two dots the trajectories seem to circulate around [Rot, p. 115-116].

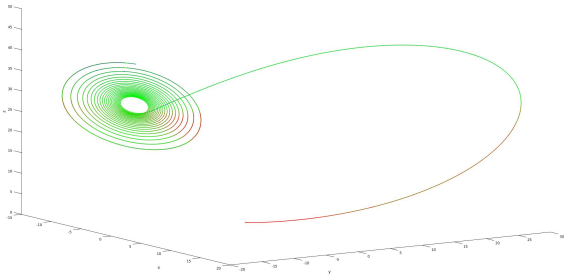
7.5.2 Different Behavior of Trajectories

Figures 7.3 and 7.4 are identical in every aspect except the timespan, which increases after each graph. In both figures the timespan the system operates is marked in the caption. For example at 5 milliseconds the trajectories have swung to the right and continued left to spiral around the left side fixed point. At 15 milliseconds neither of the trajectories have left spiraling outward around the left fixed point. At 17 millisecond the trajectories shoot back over to the right side. At this point we see how the red and the green trajectories have diverged away from each other a little. In the graphs at 19 and 21 milliseconds the trajectories keep spiraling around the left fixed point. At 22 milliseconds we can see that the red trajectory have already shot back to spiral around the right side fixed point while the green and the blue trajectories are still spiraling around the left one.

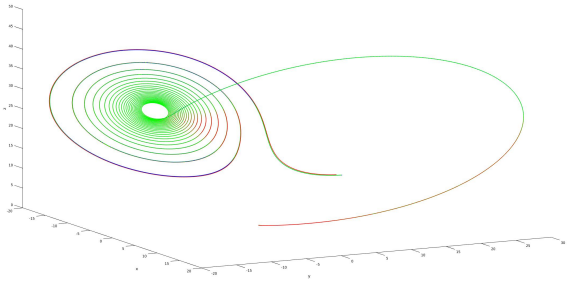
Now for the sake of an argument we treat the blue trajectory as the path we are trying to predict. Let's say that we don't have precise information of blue's initial conditions



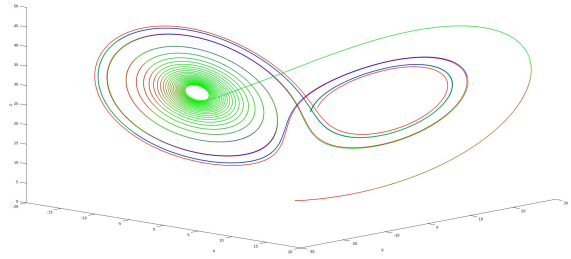
(a) 5 milliseconds.



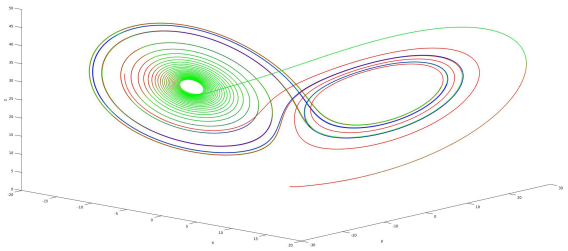
(b) 15 milliseconds.



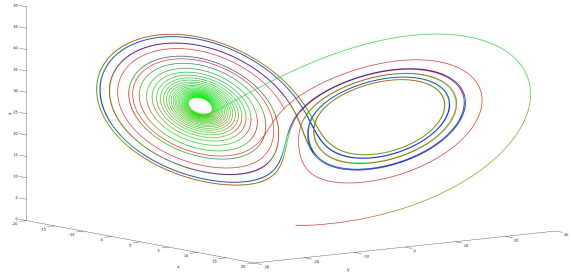
(c) 17 milliseconds.



(d) 19 milliseconds.



(e) 21 milliseconds.



(f) 22 milliseconds.

Figure 7.3: Plots of the Lorenz system with different time-spans. Each graph has three different trajectories starting from almost identical initial conditions.

$\boldsymbol{y}_B(0) = (0, 1, 0)$, but only approximations. First approximation is the green $\boldsymbol{y}_G(0) = (0, 0.99, 0)$ and the second is the red $\boldsymbol{y}_R(0) = (0, 0.9, 0)$. As can be seen in Figure 7.4 at 23 milliseconds the red and the green are closer together than is the blue's distance

from them. Of course the reader needs to trust the author's word that the viewing angle wouldn't significantly alter the perception of the distances. So future paths that were initially further result in being closer together and future paths that were initially closer together result in being further apart. Hereby is figuratively shown how prediction breaks down as increasing measurement accuracy doesn't increase the prediction accuracy.

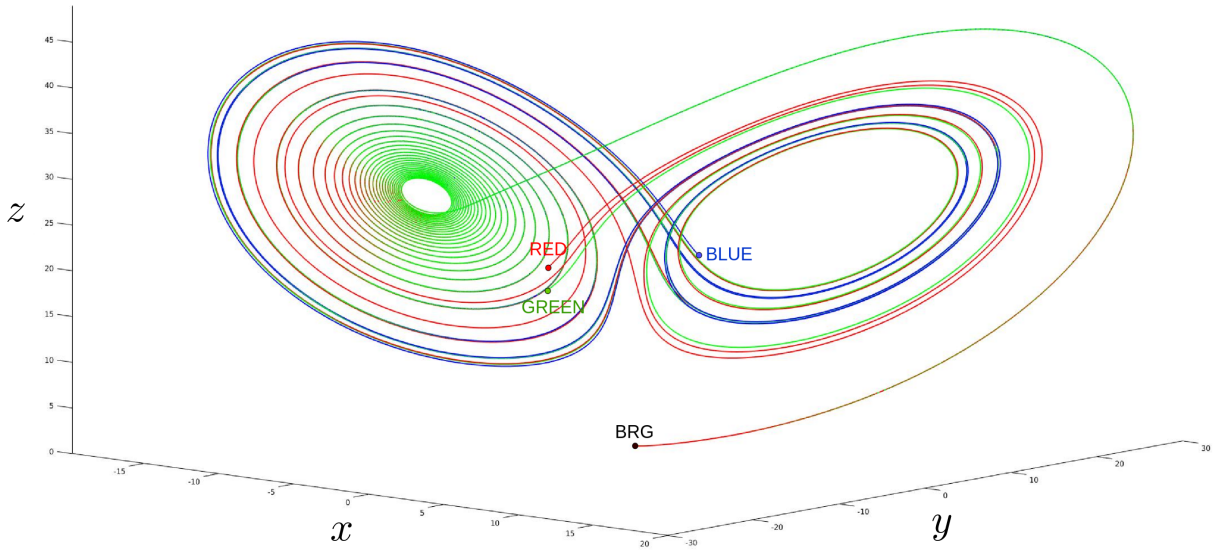


Figure 7.4: Timespan 23 milliseconds. Plots of three trajectories of the Lorenz system.

Chapter 8

Analysis of the Secret Order

8.1 Volume Contraction

The Lorenz system is *dissipative*. What this means is that volumes in phase space contract under the flow. To get an general understanding of the latter we can think of any three-dimensional system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x})$ and pick an arbitrary closed surface $S(t)$ in phase space. Next we take the points on S as initial conditions for future trajectories. When we let the trajectories evolve for an infinitesimal time dt we see the S evolving into a new surface $S(t + dt)$. The volume evolution we are concerned of is the volume what is left between S and $S(t + dt)$.

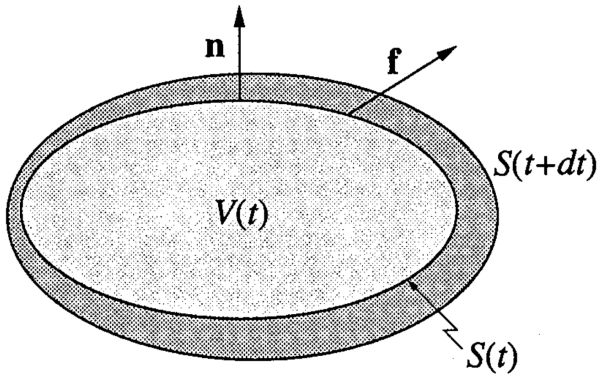


Figure 8.1: A side view of the volume evolution [Str94, p. 312].

More specifically expressed we denote $\boldsymbol{n}$ as the outward normal on S. Consequently $\boldsymbol{f} \cdot \boldsymbol{n}$

is the outward normal component of velocity as $\boldsymbol{f}$ is the velocity of the points. Thus up to a first order of dt the volume of $(\boldsymbol{f} \cdot \boldsymbol{n}dt)dA$ is swept out by an area dA in time dt as shown in Figure 8.2.

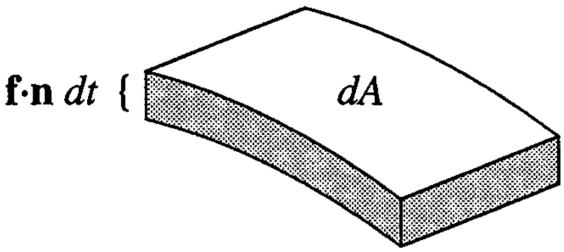


Figure 8.2: A volume of $(\boldsymbol{f} \cdot \boldsymbol{n}dt)dA$ evolves in time dt [Str94, p. 312].

When we denote V_* as the volume swept out by tiny patches of surface, integrated over all patches, we can write

$$V(t + dt) = V(t) + V_*.$$

So we obtain

$$V(t + dt) = V(t) + \int_S (\boldsymbol{f} \cdot \boldsymbol{n}dt)dA,$$

yielding

$$\dot{V} = \frac{V(t + dt) - V(t)}{dt} = \int_S \boldsymbol{f} \cdot \boldsymbol{n}dA.$$

Lastly, we rewrite the integral above in accordance with the Divergence Theorem presented briefly in Appendix A.5, and get

$$\dot{V} = \int_V \nabla \cdot \boldsymbol{f}dV. \tag{8.1}$$

For the Lorenz system we have

$$\boldsymbol{f}(\boldsymbol{y}) = \begin{pmatrix} a(y - x) \\ x(c - z) - y \\ xy - bz \end{pmatrix}, \tag{8.2}$$

so

$$\begin{aligned}\nabla \cdot \boldsymbol{f} &= \frac{\partial}{\partial x}[a(y-x)] + \frac{\partial}{\partial y}[cx - y - xz] + \frac{\partial}{\partial z}[xy - bz] \\ &= -a - 1 - b \\ &= -10 - 1 - 28 \\ &< 0.\end{aligned} \tag{8.3}$$

Since the divergence is constant, (8.1) reduces to

$$\dot{V} = -(a + 1 + b)V. \tag{8.4}$$

(8.4) has solution $V(t) = V(0)e^{-(a+1+b)t}$. Now we can say that when regarding the Lorenz system, volumes in phase space shrink exponentially fast.

What this means is that any arbitrary massive solid glob of initial conditions eventually shrinks to a limiting set of zero volume. Metaphorically one could picture a balloon with the air being sucked out of it. All trajectories initiating in the glob wind up somewhere in this limiting set. In case of the Lorenz system we see that the limiting set consists of fixed points around which the trajectories spiral around drawing a large amount of surfaces. The number of circuits made around each fixed point varies from one cycle to the next in a very unpredictable fashion. Indeed, the sequence of the number of circuits on each side displays a large amount of the characteristics of a *random* sequence. When viewing the trajectory of the Lorez system, see Figure 6.4, it appears to settle onto a supremely thin set that looks, by chance, like a pair of butterfly wings. This limiting set is the attracting set of zero volume [Str94, p. 312-313].

What is the geometrical structure of this attracting set of zero volume? Figure 6.4 suggests that it is a collection of surfaces that somehow merge into each other. The Uniqueness Theorem states that this merging needs to happen in a manner in which the different parts of the trajectory don't cross [Str94, p. 319-320]. As it turns out it is not simple to describe geometrical features of attractors as even defining such entities seems to cause trouble.

8.2 Defining Attractor

Defining an attractor is far from a simple task as a definition that is broad enough to include all the natural candidates, but restrictive enough to exclude the impostors would

naturally be preferred. Differing opinions about what the exact definition should be make completion of the definition complicated. When speaking loosely, one could define an attractor to be a set to which all neighboring trajectories converge. Stable limit cycles and stable points are excellent examples of such sets. More precisely, we define an attractor in a following manner.

Definition 8.2.1. An attractor is a closed set X with the following properties:

1. X is an invariant set. Any trajectory $\boldsymbol{y}(t)$ starting in X will stay in X when $t \to \infty$.

2. There exists an open set U containing X such that if $\boldsymbol{y}(0) \in U$, then the distance from $\boldsymbol{y}(t)$ to X tends to zero as $t \to \infty$. In other words X attracts an open set of initial conditions meaning that X attracts all trajectories that start sufficiently close to it. The largest U from which trajectories are attracted to X is called the basin of attraction of X.

3. X has not proper subset that satisfies conditions 1. and 2. In other words X is minimal.

Despite the fact that all trajectories of the Lorenz system wind up to a bounded set of zero volume it might just be that that set is not an attractor as it might not be minimal. In this rather technical sense no one has been able to show that the Lorenz attractor seen in computer experiments is truly an attractor. Still, a significant majority believes that it is [Str94, p. 324].

8.3 Exponential Divergence of Nearby Trajectories

As shown in Chapter 7 the motion of the trajectories on the attractor displays sensitive dependence on initial conditions. Correspondingly two trajectories with almost identical initial conditions will have totally different futures as they diverge rapidly from each other. Supposing we measure the initial conditions of two trajectories within Lorenz system very accurately, we can show that there is a time horizon beyond which prediction breaks down. In Section 7.5 of Chapter 7 it is demonstrated how this time horizon for the Lorenz system is no longer than 23 milliseconds when measuring at an accuracy of two decimals. This is shown schematically in Figure 8.3 as trajectories with almost identical initial conditions diverge significantly quickly [Str94, p. 320-322].

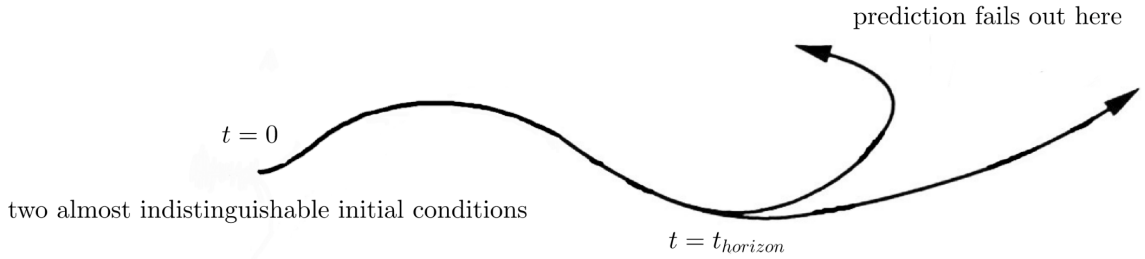


Figure 8.3: Schematic of a time horizon $t_{horizon}$ after which prediction breaks down.

8.4 Strange Attractor

Finally, we get to the topic of *strange attractor*. When examining the Figure 6.4 in detail one realizes that the figure consists of "infinite complex of surfaces". Today such a complex would be called a fractal. An example of a fractal is a set of points with zero volume but infinite surface area. The name *strange attractor* derives from the fractal like nature of attractors as they are often fractal sets. Since in the last chapters we have tried to study the concept of sensitive dependence on initial conditions we define a strange attractor as an attractor exhibiting such behavior. Today the dynamical property of sensitive dependence on initial conditions is regarded as the more important property than the geometric property of fractal nature but this is truly a question of preference. The term *chaotic attractor* is used when one wishes to emphasize the dynamical properties of a system and the term *fractal attractor* when the geometrical properties are more of interest [Str94, p. 323-325].

8.5 Explaining Chaos

When explaining chaos, no particular book or work is above the others but Peitgen's, Jürgens' and Saupe's work *Chaos and fractals - new frontiers of science* [PJS04] must be mentioned again as it is a very comprehensive survey on the topic. In similar fashion no conclusive universally accepted definition for the term *chaos* exists. Nonetheless Strogatz's easily approachable oeuvre about nonlinear dynamics and chaos [Str94, p. 323-324] presents a working definition:

> *"Chaos is aperiodic long-term behavior in a deterministic system that exhibits sensitive dependence on initial conditions."*

In essence Strogatz's working definition can be seen as framed by three widely accepted elements. A short list of these three elements certainly does no justice when describing something as wide and complex as *chaos*. Nevertheless, such a list is still presented here to conclude this thesis.

I. "Chaotic system must display aperiodic long-term behavior" meaning that trajectories which do not settle down to fixed points or periodic orbits exist as $t \to \infty$. For example chaotic system could demand that aperiodic trajectories should occur with nonzero probability, given a random initial condition.

II. "Chaotic system must be deterministic" meaning that the system involves no random factors. A chaotic system has no random or noisy inputs or parameters. Nonlinearity is the cause of any irregular behaviour the system displays, not noisy driving forces.

III. "A chaotic system exhibits sensitive dependence on initial conditions."

Appendix A

A.1 Existence Theorem and Uniqueness Theorem

To be able to plausibly discuss solutions of Lorenz system we present two important theorems regarding initial value problems of nonlinear differential equations and their solutions.

Theorem A.1.1 (Existence Theorem)**.** *If f is continuous on an open box*

$$R : a < x < b, \quad c < y < d$$

that contains (x_0, y_0) then the following

$$y' = f(x, y), \quad y(x_0) = y_0 \tag{A.1}$$

has at least one solution on some open subinterval of (a, b) that contains x_0.

Theorem A.1.2 (Uniqueness Theorem)**.** *If both f and f_y are continuous on R then (A.1) has a unique solution on some open subinterval of (a, b) that contains x_0.*

A.2 General Solution for First-Order Differential Equation

Generally first-order differential equations such as (7.7) are seen as *initial value problems.* In the following we present a general initial value problem solving method.

Consider a general case consisting of differential equation

$$\frac{dy}{dt} = ay - b \tag{A.2}$$

and the initial condition

$$y(0) = y_0, \tag{A.3}$$

where y_0 is an arbitrary initial value. If $a \neq 0$ and $y \neq \frac{b}{a}$ we can rewrite (A.2) as

$$\frac{\frac{dy}{dt}}{y - \frac{b}{a}} = a \tag{A.4}$$

By integrating both sides we find that

$$\ln \left| y - \frac{b}{a} \right| = at + C, \tag{A.5}$$

where C is an arbitrary integration constant. Then, taking the exponential of both sides of (A.5) and solving for y, we obtain

$$y = \frac{b}{a} + ce^{at}, \tag{A.6}$$

where $c = \pm e^C$ is also arbitrary.

Finally, the initial condition (A.2) requires that $c = y_0 - \frac{b}{a}$, so the solution of the initial value problem (A.2) is

$$y = \frac{b}{a} + [y_0 - \frac{b}{a}]e^{at}, \tag{A.7}$$

reducing to

$$y = y_0 e^{at}, \tag{A.8}$$

when $b = 0$ [WR12, p. 11]. Here we note that a solution can be found also when the initial condition is $y_0 = \frac{b}{a}$

$$y = \frac{b}{a} + [\frac{b}{a} - \frac{b}{a}]e^{at} \tag{A.9}$$

$$= \frac{b}{a}. \tag{A.10}$$

According to Theorem (A.1.2) besides being a general solution, (7.8) is also the only solution.

A.3 Cauchy-Schwarz Inequality

Lemma A.3.1. *If X is a real inner product space, then for all $x, y \in X$ we have*

$$|(x, y)| \leq ||x|| \times ||y||, \tag{A.11}$$

in which equality holds if and only if the vectors x and y are linearly dependent.

Proof. Let $x, y \in X$ be nonzero and define the quadratic function $Q : \mathbb{R} \to \mathbb{R}$ by

$$Q(\lambda) = ||x + \lambda y||^2 = ||y||^2\lambda^2 + 2(x, y)\lambda + ||x||^2. \tag{A.12}$$

Clearly, $Q(\lambda) \geq 0$ for each $\lambda \in \mathbb{R}$. So the discriminant of the quadratic function is non-positive, that is, $4|(x, y)|^2 - 4||x||^2 \leq 0$ or $|(x, y)| \leq ||x|| \times ||y||$. Equality holds if and only if the quadratic has a zero, that is, if and only if there exists some $\lambda \in \mathbb{R}$ such that $x+\lambda y = 0$, which is, of course, equivalent to having the vectors x and y linearly dependent [AB03, p. 246]. □

A.4 Complex Numbers

Here we show that we can ignore the imaginary part of the right hand side of (7.13). To do that we require some basics of complex numbers.

Basic Properties of Complex Numbers

We start by viewing a complex number $z = x + iy \in \mathbb{C}$ representing a point or a vector (x, y) in $\mathbb{R}^2$. A complex number has a real part $x = Re(z)$ and an *imaginary part* $y = Im(z)$ and it has an absolute value, also known as modulus, $r = |z| = \sqrt{x^2 + y^2}$.

Exponential and Trigonometric Functions

We have

$$exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \ldots, \quad z \in \mathbb{C}. \tag{A.13}$$

Hence we have also

$$\frac{d}{dz} exp(z) = exp(z), \quad z \in \mathbb{C}. \tag{A.14}$$

As (A.14) yields to a fundamental property of

$$exp(z_1 + z_2) = exp(z_1)exp(z_2), \quad z_1, z_2 \in \mathbb{C},$$

we see that

$$1 = exp(0) = exp(z - z) = exp(z)exp(-z),$$

which essentially says

$$exp(-z) = \frac{1}{exp(z)}.$$

When considering (A.13) we see that

$$\overline{\sum_{n=0}^{\infty} \frac{z^n}{n!}} = \sum_{n=0}^{\infty} \frac{\overline{z}^n}{n!}.$$

Thus we have $\overline{exp(z)} = exp(\overline{z}), \quad z \in \mathbb{C}$ [Ber12, p. 7-26].

Ignoring the Imaginary Part

Now we can derive the property we initially needed for (7.13)

$$\begin{aligned} |exp(z)|^2 &= |\overline{exp(z)}exp(z)| \\ &= |exp(\overline{z})exp(z)| \\ &= |exp(\overline{z} + z)| \\ &= |exp\left(2Re(z)\right)| \end{aligned}$$

In summary we have

$$|e^z|^2 = |e^{2Re(z)}|,$$

which essentially says

$$|e^z| = e^{Re(z)}.$$

A.5 Divergence Theorem

The Divergence Theorem of Gauss states that if $\boldsymbol{f}$ is a vector function of position with continuous derivatives, and V is the volume bounded by a closed surface S, then

$$\iiint_V \nabla \cdot \boldsymbol{f} dV = \iint_S \boldsymbol{f} \cdot \boldsymbol{n} dS, \tag{A.15}$$

where n is the positive (outward drawn) normal to S and divergence is defined as $\nabla \cdot \boldsymbol{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$ [Mur59, p. 106].

Appendix B

Matlab Code

B.1 Lorenz System Plot

```
clear all, close all, clc


x=-20:0.1:20;
y=-20:0.1:20;
z=-20:0.1:20;


Beta = [10; 28; 8/3]; %chaotic parameter values
x0 = [0; 1; 0]; %initial condition
dt = 0.001;
tspan = dt:dt:50;

options = odeset('RelTol',1e-12,'AbsTol',1e-12*ones(1,3));
[t,x] = ode45(@(t,x)lorenz(t,x,Beta),tspan,x0,options);




plot3(x(:,1),x(:,2),x(:,3),'r','Linewidth',1.5);
ax=gca
set(gca,'color','w','xcolor','k','ycolor','k','zcolor','k');
```

```
set(gcf,'color','w');
```

B.2 Prediction Breaks Down: Three Trajectories

```
clear all, close all, clc


x=-20:0.1:20;
y=-20:0.1:20;
z=-20:0.1:20;


Beta = [10; 28; 8/3]; %chaotic parameter values
x0 = [0; 1; 0]; %initial condition
dt = 0.001;
tspan = dt:dt:23;

options = odeset('RelTol',1e-12,'AbsTol',1e-12*ones(1,3));
[t,x] = ode45(@(t,x)lorenz(t,x,Beta),tspan,x0,options);




plot3(x(:,1),x(:,2),x(:,3),'b','Linewidth',1.5);
xlabel('x');
ylabel('y');
zlabel('z');
ax=gca
set(gca,'color','w','xcolor','k','ycolor','k','zcolor','k');
set(gcf,'color','w');


hold on;
Beta = [10; 28; 8/3]; %chaotic parameter values
x0 = [0; 0.9; 0]; %initial condition
dt = 0.001;
tspan = dt:dt:23;
```

```
options = odeset('RelTol',1e-12,'AbsTol',1e-12*ones(1,3));
[t,x] = ode45(@(t,x)lorenz2(t,x,Beta),tspan,x0,options);




plot3(x(:,1),x(:,2),x(:,3),'r','Linewidth',1.5);
xlabel('x');
ylabel('y');
zlabel('z');
ax=gca;



hold on;
Beta = [10; 28; 8/3]; %chaotic parameter values
x0 = [0; 0.99; 0]; %initial condition
dt = 0.001;
tspan = dt:dt:23;

options = odeset('RelTol',1e-12,'AbsTol',1e-12*ones(1,3));
[t,x] = ode45(@(t,x)lorenz2(t,x,Beta),tspan,x0,options);




plot3(x(:,1),x(:,2),x(:,3),'g','Linewidth',1.5);
xlabel('x');
ylabel('y');
zlabel('z');
ax=gca;

hold off;
```

B.3 Distribution of Maximum Jacobian Matrix's Eigenvalues

```
x=-1000:10:1000;
y=-1000:10:1000;
z=-1000:10:1000;


max_real_eig_xyz = zeros(length(x),length(y),length(z));

for k=1:length(z)



    for l=1:length(y)

        for m=1:length(x)

            J = [-10 10 0; 28-z(k) -1 -x(m); y(l) x(m) -8/3];
            max_real_eig_xyz(m,l,k)=max(real(eig(J)));




        end

    end
end



max_eig_xyz = zeros(length(x),length(y));

for k=1:length(x)

    for l=1:length(y)
```

```
            max_eig_xyz(l,k) = max(max_real_eig_xyz(l,k,:));


        end
end

[X,Y]=meshgrid(x,y);
figure;


color = zeros(size(max_eig_xyz));
color(max_eig_xyz < 0) = -1;
color(max_eig_xyz >= 0) = 1;
surf(X,Y,max_eig_xyz,color);
xlabel('x');
ylabel('y');
zlabel('max {\lambda}');
```

Bibliography

[3DP] Customized 3D Plotter, Calcplot3d. http://www.monroecc.edu/faculty/paulseeburger/calcnsf/CalcPlot3D/. 2019-05-08.

[AB03] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer-Verlag Berlin Heidelberg, 2003.

[Ber12] Christian Berg. Complex Analysis, Online Notes. http://web.math.ku.dk/noter/filer/koman-12.pdf, 2012. 2019-08-08.

[BT09] H. Broer and F. Takens. *Dynamical Systems and Chaos*. Applied Mathematical Sciences. Springer, New York, 2009.

[Har03] Michael D. Hartl. Lyapunov exponents in constrained and unconstrained ordinary differential equations. Technical report, Department of Physics, California Institute of Technology, Pasadena CA 91125, 5 2003.

[HDS74] Morris W. Hirsch, Robert L. Devaney, and Stephen Smale. *Differential Equations, Dynamical Systems, and Linear Algebra Volume 60 of Pure and Applied Mathematics Pure and Applied Mathematics, a Series of Monographs and Tex*. Applied Mathematical Sciences. Academic Press, 1974.

[Hol] G. B. Holland. Higher-Order Derivatives and Taylor's Formula in Several Variables. https://sites.math.washington.edu/~folland/Math425/taylor2.pdf. Unpublished paper, 2019-15-08.

[Joh08] Aylmer Johnson. A Survival Guide to Vector Calculus. http://www2.eng.cam.ac.uk/~alj3/vc.pdf, 2008. Online notes, Unpublished paper, 2019-08-05.

[Jun69] Carl Jung. *Archetypes and the Collective Unconscious, Collected Works of C.G. Jung, Volume 9, Part 1*. Princeton, N.J.: Princeton University Press, 1969.

[Kau11] Richard Kautz. *Chaos : The Science of Predictable Random Motion*. Oxford University Press, New York, 2011.

[Kel94] Stephen H. Kellert. *In the Wake of Chaos : Unpredictable Order in Dynamical Systems*. The University of Chicago Press, United States of America, 1994.

[Mur59] Spiegel R. Murray. *Vector Analysis and an Introduction to Tensor Analysis*. Schaum Pub. Co., New York, 1959.

[PJS04] Heinz-Otto Peitgen, Hartmut Jürgens, and Dietmar Saupe. *Chaos and Fractals - New Frontiers of Science (2. ed.)*. Springer, 2004.

[Rot] Daniel Rothman. Nonlinear Dynamics I: Chaos, Fall 2006. `https://dspace.mit.edu/bitstream/handle/1721.1/84612/12-006j-fall-2006/contents/lecture-notes/lecnotes11.pdf`. Online notes, Unpublished paper, 2019-07-10.

[Rud76] Walter Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.

[Spa12] C. Sparrow. *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*. Applied Mathematical Sciences. Springer New York, 2012.

[Str94] S.H. Strogatz. *Nonlinear Dynamics and Chaos With Applications to Physics, Biology, Chemistry, and Engineering*. Studies in nonlinearity. Wesview Press, Perseus Books, 1994.

[Wig03] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Texts in Applied Mathematics. Springer New York, 2003.

[WR12] Boyce E. William and DiPrima C. Richarnd. *Elementary Differential Equations and Boundary Value Problems*. 2012.